HYBRID FOREST MODELLING OF *Pinus radiata* D. DON IN CANTERBURY, NEW ZEALAND

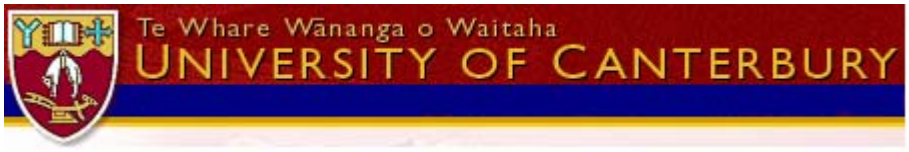
A thesis
Submitted in partial fulfilment
Of the requirement for the degree of
Doctor of Philosophy in Forestry
In the University of Canterbury

by

Guy L. Pinjuv

University of Canterbury

2006



CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

During this study two models were developed to predict growth of *Pinus radiata* D.Don plantations in Canterbury, New Zealand. The first, CanSPBL(1.2), is a model for whole rotations of stands owned by Selwyn Plantation Limited in Canterbury. The second model, CanSPBL(water) is a hybrid growth model for the Selwyn estate in Canterbury that incorporates an index of root zone water balance over the simulation period. An existing stand growth and yield model CanSPBL was examined using a validation dataset of PSP measurements that were not used in model fitting. Projection bias was shown for mean top height, basal area per hectare, and residual stand stocking particularly for stands at elevations exceeding 450 metres.

The new model, CanSPBL(1.2) showed an increase in precision of 4 – 46% over CanSPBL(1.0) at a stand level. The components of the stand model include mean top height, basal area per hectare, stems per hectare, and diameter distribution. The mortality model was made in conjunction with managers at CanSPBL to exclude catastrophic mortality events from model projections. Data used for model fitting was filtered using a mortality index based on the -3/2 power law. An examination of this model with an independent dataset showed little apparent bias.

The new model, CanSPBL(water) was developed to include an index of water balance over the simulation period. Water balance estimates were made using a sub model for root zone water balance included in the hybrid physiological model 3-PG (Landsberg and Waring, 1997). The new model showed an increase in precision of 1 – 4% over CanSPBL(1.2) at a stand level (with the exception of the model for maximum diameter which showed a decrease in precision of 0.78%) using climatic inputs that included yearly variation. However the model showed increases of precision from 0.5 to 8% (with the exception of maximum diameter again, showing a decrease in precision of 0.13%) using long term monthly average climatic inputs. The components of the stand model also include mean top height, basal area per hectare, stems per hectare, and diameter distribution. The mortality model was also fitted with a data set filtered using a mortality

severity index based on the -3/2 power law to exclude catastrophic mortality events. An examination of this model with an independent dataset showed little apparent bias.

Two models to predict a one sided canopy leaf area index (LAI) of radiata pine stands in the Canterbury Plains of New Zealand were also developed. The models were fitted using non-linear least squares regression of LAI estimates against stem measurements and stand characteristics. LAI estimates were derived from digital analysis of fisheye lens photography. The models were kept simple to avoid computational circularity for physiological modelling applications.

This study included an objective comparison and validation of a range of model types. The models CANTY (Goulding, 1995), CanSPBL(1.2) (Pinjuv, 2005), CanSPBL-water (Pinjuv, 2005), and 3-PG (Landsberg and Waring, 1997) were compared and validated with the main criteria for comparison being each model's ability to match actual historical measurements of forest growth in an independent data set. Overall, the models CanSPBL(water), and CanSPBL(1.2) performed the best in terms of basal area and mean top height prediction. Both models CanSPBL(water), and CanSPBL(1.2) showed a slightly worse fit in predictions of stocking than did the model CANTY. The hybrid model 3PG showed a better fit for the prediction of basal area than the statistically based model CANTY, but showed a worse fit for the prediction of final stocking than all other models. In terms of distribution of residuals, CanSPBL(1.2) had overall the lowest skewness, kurtosis, and all model parameters tested significant for normality. 3PG performed the worst on average, in terms of the distribution of residuals, and all models tested positively for the normality of residual distribution.

CHAPTER 1

INTRODUCTION

1.1 Background

Modern forest resource managers face challenging decisions in trying to provide forest products for a growing world population despite a shrinking natural resource base challenged by global climate change, desertification, environmental pollution, and other stresses. Managers also need to ensure that forest products are provided to society in ways that are both ecologically and economically sustainable (Peng, 2000b).

Sustainable forest management hinges on accurate estimates of future growth of the existing resource. Forest managers require simulation models to predict the potential impacts of future changes in the global environment. Using models, forest managers can examine the consequences of alternative actions without incurring major costs or losses through repeated trial and error.

Prediction is critical for decision-making, and the quality of a decision is strongly influenced by the quality and types of models used. Classical growth and yield forest modelling has been criticised as being empirical, with any derived relationships of growth revealing little about physiological control mechanisms or being able to adapt to changing environmental conditions. On the other side of the modelling spectrum, complex mechanistic models of growth have been criticized as being cumbersome, requiring many hard to measure inputs and relying heavily on inadequately tested assumptions. There is now a continuum from the most empirical of forest growth models that fit arbitrary regression equations to growth data, through to complex mechanistic models, with hybrid mixtures of the two approaches in between. The point in this spectrum that will be most appropriate to answering questions related to forest management will depend on the questions being asked, the scale of those questions, and the availability of input data.

This study covers a broad research area in forest modelling systems. The main goal of the project is to evaluate a range of stand-level modelling approaches, from traditional growth and yield systems to models that explicitly represent environmental influences, photosynthesis and carbon allocation, such as the 3-PG model (Landsberg and Waring, 1997). Models were fitted using data from a specific region managed by a forest company, and their performance was assessed when they were applied to an independent dataset. The study involved some biomass and canopy measurements that provided the basis for a sub-model to predict canopy leaf area index that was used as an input into other models. This new direction in growth and yield modelling will hopefully help bridge the gap between statistical and physiological approaches to forest modelling and give some insight as to what level of resolution is appropriate for answering questions related to forest management.

1.2 Research Objectives

The intent of the research is to model growth and yield and compare forest modelling systems for plantation-grown *Pinus radiata* D. Don on Selwyn Plantation Board Ltd. (SBPL) land. The following is a list of study objectives that were set out based on the existing state of the art in modelling research approaches, the requirements of the forestry company SPBL, and research already completed in growth and yield modelling of the SPBL estate by Zhao (1999).

(i) Test existing model CanSPBL with new data.

(ii) Update existing growth and yield model on SBPL land (Zhao 1999) with new sample data. Specifically update growth and yield estimates on SBPL forested sites in hill country.

(iii) Compare alternative strategies for mixing conventional growth and yield models with those that include a higher level of biological detail (usually called "process-level models"). The accuracy of any given mixture of models to project independent measures of tree and stand growth in permanent sample plots throughout the SPBL estate will be the main

criterion for choosing a modelling strategy. The following approaches will be taken to achieve this aim;

(a) Construct leaf area index (LAI) model for SBPL lands. This model will be used to attach LAI information to permanent sample plot data in the current and past inventories. LAI predictions will be used as an input into other models to predict physiological information for all plots.

(b) Model water balance over the simulation period using existing models (Landsberg and Waring 1999, Whitehead *et al.*, 2001; Watt *et al.* 2003).

(c) Incorporate water balance into updated growth and yield model. This will be the first physiological term added to this model and its performance will be analysed via residual analysis.

(d) Calibrate existing hybrid model, 3-PG (Landsberg and Waring, 1997) to fit growth measurements of forest stands on SBPL land. This will be done using local parameters for radiata pine growth and leaf area predictions made by developed LAI model. Model outputs at a stand level and model performance will be compared to traditional growth and yield models with and without physiological terms.

1.3 Symbols

The following is a list of symbols and definitions used throughout this thesis. These will be the assumed meaning of all symbols and definitions unless otherwise stated.

α, β, γ or *a,b,c:*	parameter of regression coefficients in an equation
a:	1x3 matrix in context of CANTY model
A:	3x3 matrix
Alt:	altitude
AMB:	average model bias
APAR:	absorbed photosynthetically active radiation (nm)
ASW:	available soil water deficit as estimated from water balance models (mm), or available soil water input into 3-PG water balance model (mm) depending on context
b:	1x3 matrix in context of CANTY model
b_h:	coefficient in CANTY site index equation
$b_h t$:	scaled time
BA:	net basal area per hectare (sectional area at breast height, m^2/ha)
B_1:	BIOCLIM climate estimate for Christchurch airport
B_2:	BIOCLIM climate estimate for a given site
B_{2adj}:	adjusted BIOCLIM climate estimate for a given site
C or *C*:	3x3 matrix in context of CANTY model, and constant in all other contexts.
CanSPBL(1.0)	growth and yield model established by Zhao (1999)
CanSPBL(1.2)	growth and yield model established by Pinjuv (2005)
CanSPBL(water)	hybrid growth and yield model established by Pinjuv (2005)
CanSPBL(water_using average climate)	refers to results from running hybrid growth and yield model CanSPBL(water) established by Pinjuv (2005) with average climatic inputs
CO_2:	carbon dioxide
c_θ:	parameter that describes different soil types in soil water modifier
d, dbh or dbhob:	individual tree diameter at breast height outside bark (cm) or *d* = distance depending on context

D:	diameter class (cm), or drainage from the soil surface (mm) depending on context
$\overline{d}$:	arithmetic mean diameter (cm)
$d_{\bar{g}}$:	diameter of mean basal area
d_{max} ,or *DMAX*:	maximum diameter at breast height in a plot or stand (cm)
d_{min}:	minimum diameter at breast height in a plot or stand (cm)
d_{pi}:	i^{th} dbh in the p^{th} plot (cm)
d_{std} ,or *DSTD:*	standard deviation of diameter in a plot or stand (cm)
d_{var}:	variance of diameter in a plot or stand (cm)
E:	evapotraspiration (mm)
EF:	model efficiency
elev:	altitude above sea level (m)
E_{bi}:	transpiration from the broom canopy (mm)
$E_{bwi:}$	evaporation from the wet broom canopy (mm)
E_{gi}:	evaporation from the soil surface (mm)
E_{ti}:	transpiration from the tree canopy (mm)
$E_{twi:}$	evaporation from the wet tree canopy (mm)
ε :	radiation use efficiency
F_a:	relative age of a forest or plantation
f_{age}:	age modifier
f_{N}:	soil nutrition modifier
f_{sw} or f_{θ} :	soil water modifier
f_{T}:	temperature modifier
f_{VPD} or f_D :	vapour pressure deficit modifier
F_i:	drainage from the root zone (mm)
G:	net basal area per hectare (sectional area at breast height, m^2/ha)
g_c:	canopy conductance (m s^{-1})
$g_{c_{\max}}$:	maximum canopy conductance (m s^{-1})
GPP:	gross primary productivity
h:	total individual-tree height from ground level (m)
H, or MTH:	mean top height (m)
ha:	hectare
I:	canopy interception (mm)

k:	light extinction coefficient, or relationship between stomatal conductance and vapour pressure deficit (depending on context)
LAI:	one sided leaf area index estimated from an optical measuring device from equation
LAI_{ds}:	one sided leaf area index determined from destructive sampling
M:	potential plots
m:	sample size
m_1:	climate measurement at Christchurch airport
MSE, or RMS:	mean square error of regression
N:	number of stems per hectare or stocking (stems/ha) or nitrogen depending on context
N_{adj}:	adjusted number of live stems per hectare (stems/ha)
n_{age}:	parameter that controls the rate of change of f_{age} modifier
n:	total number of observations in a plot or sample size
$\overline{n}$:	average number of observations in a plot or sample size
n_p:	number of trees in the p^{th} plot
n_θ :	parameter that describes different soil types in soil water modifier
P:	precipitation (mm)
PAI:	plant area index
PAR:	photosynthetically active radiation (nm)
pFS2:	3-PG parameter for stem partitioning at diameter = 2 cm
pFS20:	3-PG parameter for stem partitioning at diameter = 20 cm
P(death):	probability of stem mortality
PSP:	permanent sample plot
P_i:	rainfall (cm)
Q_i:	ceptometer reading underneath a canopy which measures photosynthetically active radiation transmitted through the canopy
Q_0:	ceptometer reading in an open area which measures total photosynthetically active radiation
RSS, or SSE:	residual sum of squares or sums of squares error
r_θ :	moisture ratio
S:	survival of trees in a plot
SI:	site index – mean top height at 20 years of age

SPBL:	Selwyn Plantation Board Ltd.
T:	stand age (years)
T_a:	mean air temperature (°C)
T_x:	maximum air temperature (°C)
T_n:	minimum air temperature (°C)
U:	input management options into CANTY growth and yield model
v:	individual tree-stem volume (m^3)
V:	stand stem volume per hectare (m^3/ha)
VPD:	saturation vapour pressure (kPa)
W:	plant weight (kg)
W_i:	root zone water balance (mm)
wSx1000:	3-PG parameter for the maximum stem mass per tree for a stand density of 1000 trees / ha (kg / tree)
X:	Variable in an equation or dummy variable for altitude (depending on context). When used as a dummy variable $X = 0$ when altitude < 250m, and $X = 1$ when altitude ≥ 250m
Y, Y_2, or Y_1:	Variables in equations, usually used to display difference equation forms
Y_i:	observed data
$\hat{Y}_i$:	modelled data
$\overline{Y}_i$:	mean of the observed data
3-PG:	hybrid growth model established by (Landsberg and Waring 1997)
ηf :	3-PG foliage partitioning coefficient
ηs :	3-PG stem partitioning coefficient
$\Phi_{p.a.u}$:	utilizable, absorbed photosynthetically active radiation (nm)
θ_T :	root zone water balance at time T (mm)
θ_{T-1} :	root zone water balance in the previous month (mm)

CHAPTER 2

LITERATURE REVIEW

2.1 Modelling Procedures

There have been three major approaches to forest simulation modelling, classic growth and yield models (empirical), process-based models, and hybrid models (which are a mixture of both empirical and process-based models) . This chapter introduces and discusses the inherent advantages and disadvantages of using empirical and process-based models in sustainable forest management. For each modelling approach, there are weaknesses and strengths which will also be discussed further.

2.1.1 Forest growth and yield models

Growth and yield models are usually curves fitted to historical data describing forest growth on a particular site**.** Peng (2000b) calls this the "historical bioassay approach" and claims it is the simplest and most believable method of predicting future forest growth on a site given the future growing conditions and climatic variables are expected to remain constant (Kimmins, 1990; Vanclay, 1994).

Functions used to describe growth and yield are compatible in that growth is a derivative of yield. Clutter (1963) was among the first to describe growth and yield systems in terms of difference equations, where future yield is expressed as a function of existing yield and the interval in time between the two observations. Difference equations can be derived in growth and yield systems by separating yield and time variables in the derivative of the equation, then both sides of the equation are integrated in terms of T_1 and T_2, and then from Y_1 to Y_2. Desirable properties of growth and yield functions outlined by Clutter *et al.* (1983) are; (1) functions should be compatible, consistent (as T_2 approaches T_1, Y_2 should approach Y_1), (2) they should be path invariant (where predicting Y_3 from Y_1

should give the same result as predicting Y_2 from Y_1 and then Y_3 from Y_2), and (3) as T_2 nears infinity, Y_2 should approach the upper asymptote of the equation.

Sigmoid curves have been used in many growth and yield models to represent the growth of biological processes such as stand basal area or mean top height of individual trees and whole stands over time (Causton, 1983). Sigmoid forms that have been proposed for modelling plant biomass or tree size include the log-reciprocal Schumacher model (Schumacher, 1939; Clutter, 1963), the Chapman-Richards model (Von Bertalanffy, 1949; Richards, 1959; Pienaar and Turnbull, 1973), the Wiebull model (Yang *et al.,*1978), the Hossfeld (Woollons *et al*., 1990), and many others. The use of sigmoid curves in growth and yield models is a result of the consideration of growth relative to size of an organism or population (Mason, 1992). Mason (1992) asserts that if growth were a simple function of size, then the yield curve used to describe growth would be exponential. A sigmoid form of a growth curve proposed by Von Bertalanffy (1957), was derived by considering growth as the difference between anabolic and catabolic processes. Anabolic processes of growth have been described as being some function of an organism's surface area, with catabolic processes being some function of the organism's volume. For animal growth, the hypothesis of anabolic and catabolic growth can be justified, and is proposed as a justification for the use of sigmoid curves in describing forest growth and yield models. This justification does not entirely work since many of the most valuable parts of trees, especially the heartwood, contribute little to catabolism (Mason, 1992). Sigmoid curves have generally been used because they fit measured data well but cannot be physiologically justified via Von Bertalanffy's hypothesis.

A major strength of growth and yield models is that they describe a relationship between predictions of growth using a mathematical function and can be tested rigorously through residual analysis. A major flaw of growth and yield modelling according to Peng (2000a) is that it relies on historical data to predict future development, and assumes that conditions determining growth will be the same as those of the past. This may be inappropriate for large time scales with changing environmental factors such as increased temperatures as a result of global climate change or changes in soil nutrient status. Landsberg (2003), Vanclay (1994), and Kimmins *et al.* (1990) stress the need for more physiological approaches to be used as the functional relationship in these models.

2.1.2 Process-based models

Process-based forest models are designed to integrate energy, carbon, nutrient, and water cycles. These models estimate growth of the forest ecosystem using mathematical equations that represent underlying biological processes such as carbon, nutrient, and water cycles (Godfrey, 1983). Process-based models have the advantage that they can be more flexible than empirical relationships, and can be used to make predictions for changing ecological conditions. They can also be used to look at "what if" scenarios to indicate factors that may limit or influence plant growth at a particular site. Process-based models can also be parameterized to make predictions of plant growth on sites for which a given plant community has not historically been grown. A problem with process-based models is that they are likely to be relatively complex and require a number of parameter values that may not be readily available to forest managers (Korzukhin *et al*., 1996; Landsberg and Gower, 1997; Sands *et al*., 2000; Mäkelä *et al*., 2000). They may also require complex data inputs for estimating parameters and validation procedures, and also make predictions that are of a large enough spatial scale that they become unsuitable for answering needed questions to many forest managers. Process-based models have not been used extensively in forest management because they usually produce less accurate predictions of forest yield at a particular site than a conventional growth and yield model developed from historical data from that site (Battaglia and Sands, 1998).

In essence, a weakness of one type of model (empirical versus process-based models) is the strength of the other and vice versa. It is almost always possible to find an empirical model that provides a better fit for a given set of data, chiefly due to the constraints imposed by the assumptions of process-based models (Battaglia and Sands, 1998; Landsberg and Coops, 1999; Mäkelä *et al*., 2000; Peng 2000a,b; and Peng *et al*., 2002). Process-based models also incorporate many sub-models that describe individual physiological processes underlying plant growth, when combined together, will describe the growth of a plant community often over short increments of time. However, the greater model complexity of process-based models arising from the use of many sub-models and prediction of growth over short time increments can cause recursion and compounding of error. In a general sense, landscape scale process-based models of plantation productivity tend to have insufficient spatial resolution for stands that are a

common size for plantation forestry (5-100 ha). Statistical growth and yield models tend to be too site specific and lack the ability to make predictions under changing future environmental conditions (Woollons *et al*., 1997).

2.1.3 Hybrid models

Hybrid models or models that are a mix of process-based and empirical models can avoid the shortcomings of both approaches to some extent. Combining key elements of empirical and process approaches into a hybrid system can result in a model that predicts carbon dynamics, forest growth, and production in the short and long term (Kimmins, 1993; Battaglia *et al*., 1999; Kimmins *et al*., 1999; Peng, 2000b). These hybrid models can possibly provide information of the type and resolution required by managers and planners (Landsberg, 2003).

Hybrid models that are a mixture of both mechanistic and statistical models have been of two basic types: simplified mechanistic models, and classical growth and yield models with mechanistic terms. The simplified mechanistic model can make projections at a stand level and may use empirical methods as sub-models but the main model format is mechanistic in nature, or uses some form of carbon balance. The second type of hybrid model uses classical growth and yield methods with the addition of mechanistic predictor variables. This modelling approach still has a statistical basis and does not allocate growth based on a carbon balance framework.

Landsberg (2003) has referred to a list of hybrid models that may be of use to forest managers or industry whose target clients are at either the operational or planning levels; Mohren *et al.* (1984), and Mäkelä and Hari (1986), FORCYTE (Kimmins *et al.,* 1990), FORCYTE-11 (Kimmins *et al.,* 1999), Sievänen 1993, Sievänen and Burk (1993), Mäkelä (1997), 3-PG (Landsberg and Waring, 1997), FOREST-BGC (Running and Coughan, 1988), Running and Gower (1991), and PROMOD (Battaglia and Sands, 1997). In a general sense, all of these models contain some empirical relationships but are mechanistic in nature which makes them hard to test quantitatively.

There are a few hybrid models in the literature that are empirical in nature and include physiological terms. The advantages of such models are that they are likely to show a

close fit between predictions and measurements and may be capable of accurate predictions under changing climatic conditions. These models would also be statistically testable via residual analysis as to the quality of their predictions. Woollons *et al*. (1997) have included driving variables of mechanistic models such as mean temperature, solar radiation, rainfall, and soil type into a classical growth and yield modelling system and have shown a 10% improvement in predictions of basal area/ha over strict growth and yield curves. Snowdon *et al.* (1999) incorporated indices of annual climatic variation and photosynthesis into a growth model for *Pinus radiata* and found a significant improvement in short term prediction. They used predicted photosynthesis rates from a process-based model at a single site in the forest estate as an index for growth that was added to a Schumacher growth curve.

The hybrid model evaluated in this study is intended for use by forest industries. As such, users at that scale will be separated into managers on the ground and higher level organizational managers. Landsberg (2003) outlined model users and their needs; generally he felt forest managers who were on the ground would require models that are simple to operate and would require few parameter values. Johnsen *et al.* (2001), support this claim by stating substantial input data may limit the usefulness of a model to forest managers. The parameter values that are inputs to the model should be from readily available sources such as plot measurements of stand age and stocking, soil maps, or weather files. At higher organizational levels, managers may deal with large scale questions of wood flow and market requirements. They may be concerned with an estimate of the effect of lower annual temperatures on productivity or the effects of insects and diseases. At all organizational levels, Landsberg (2003), felt managers needed models that could account for changing environmental conditions and could be used to explore alternative management scenarios.

2.2 Leaf area index (measurement and modelling)

Leaf area has been shown to be highly correlated with productivity in a variety of ecosystems, including forests (Gholz, 1982; Waring, 1982; Webb *et al.*, 1983). As leaf area index (LAI) was used as an input into carbon balance and water balance models incorporated in this study it was critical that an accurate sub-model to predict LAI was

developed. As a basis for the creation of a model to predict LAI on the Selwyn Plantation Board estate LAI was measured on existing inventory plots. Development of such a model needed to incorporate driving variables that would predict LAI as some function of pre-measured plot variables such as stand basal area. Fownes and Harrington (1990), used a generalized form of an equation to predict individual tree Leaf Area (LA) from stem diameter.

There has been an enormous amount of physiological research that has established that net photosynthesis rates are dependent on light interception per unit leaf area and CO_2 assimilation by a plant or plant community. Monteith (1977) was the first to show net primary production (NPP) was linearly related to absorbed photosynthetically active radiation (APAR). This assimilation depends on total leaf area and the amount of light absorbed. Growth in plant biomass is the difference between carbon fixed by photosynthesis and that consumed by respiration. Plant canopies are the only resource forests have to absorb solar energy or (utilizable, absorbed photosynthetically active radiation $\Phi_{p.a.u.}$): the growth of single plants is correlated with their leaf area and stand growth is correlated with LAI, i.e.we can expect a direct relationship between production of dry mass and interception of radiation, while LAI is a major determinant of photosynthetic production (Atwel *et al.,* 1999). Wilson (1981) has shown net photosynthetic rate to be dependent on light interception on leaf area and CO_2 assimilation of intercepted light. Hunter *et al.* (1987) found annual radiata pine growth was linearly related to foliage mass and percentage of foliar nitrogen, which highlights the critical importance of nitrogen as a raw material for chlorophyll. These relationships provide some of the basis and flexibility of the analysis of assimilation and growth in terms of light interception and the structure of plant stands.

A model of forest production that is some function of LAI may need to account for the attenuation of light by foliage that will lead to lower levels of radiation available for plant growth deeper in the canopy. The Beer-Lambert equation which describes the interception of light by the canopy (with the assumption that canopies are comprised of a cloud of homogeneous randomly distributed particles), and can be written as:

$$APAR = PAR\left(1 - e^{\left(\frac{-k*LAI}{\cos(\beta)}\right)}\right) \quad (1.0)$$

Where: PAR is incoming photosynthetically active radiation above the canopy, *k* is a constant known as the light extinction coefficient, and β is the zenith angle of the sun. In practice *k* is often about 0.5 but it should vary depending on the shape and clumping of foliage within a given canopy. However, as discussed in Jarvis and Leverenz (1983) and Pierce and Running (1988), the Beer-Lambert Equation is fairly insensitive to violations of these assumptions. The non-random distribution of branches and leaves in a canopy or "clumping" of foliage violates the assumption of random distribution and introduces a bias into the Beer-Lambert estimate of APAR. Foliage clumping can occur at several levels and vary within canopies, whorls, branches, and twigs. Canopy clumping has been shown to affect estimates of APAR. Cescatti (1998) has estimated that needle clumping in crowns increases average canopy transmittance at the base of the canopy by as much as 10.9% for diffuse radiation. Whitehead *et al.,* (1990) have also quantified the effects of clumping on radiation interception by plantation-grown radiata pine. They have shown the average probability of beam penetration is greater for trees with more clumped foliage, and have estimated an index of foliage dispersion for radiata pine tree crowns. The index of foliage dispersion estimated for photosynthetically active radiation was 3.3 and 2.3 for trees with the most and least clumped foliage, respectively. For the trees included in their study, the average probability of beam penetration was greater than would have resulted if a random distribution of foliage was assumed. Canopy clumping appears to be species and age specific, and until more exact information is available on the effect of canopy clumping on APAR any use of the Beer Lambert equation will require that the assumption of random distribution of foliage particles within canopies, and the possibility of some bias in its estimate of APAR, be accepted.

LAI was used in this study as an input into other models (Watt *et al.,* 2003; Landsberg and Waring, 1997) to assess a range of modelling procedures. A prediction model for LAI is a critical component of the study. The model will be used to predict LAI for all permanent sample plots on the Selwyn estate, which will be developed from a sample of PSP's that were measured in the 2004 summer inventory. Development of a sub-model

for leaf area will require some estimate of LAI. The following is a review of some possible methods that can be used to estimate LAI.

2.2.1 Measuring Leaf Area using LI-COR, LAI-2000

One method available to measure LAI is the use of a canopy analyser (LAI- 2000, LI-COR Inc., Lincoln, NE, USA). This instrument is a portable integrating radiometer that provides relatively quick and non-destructive means of measuring LAI. The device measures sunlight intercepted by the canopy of a forest or individual tree. Such measures estimate the plant area index (PAI) since they take into account the light intercepted by the leaves, branches, and the fruiting bodies of the tree. PAI measured using the LAI-2000 is usually lower than LAI, which is probably due to deviations from four theoretical assumptions used in calculation of PAI (LI-COR-1991), outlined in detail by Cherry *et al.* (1998). The most critical theoretical assumption made is that no radiation is reflected or transmitted by the foliage. Light sensors on the device make up for this in part by only measuring light of frequencies greater than 490 nm: there is relatively little transmission or reflection of radiation in those frequencies by the foliage. The second assumption is that foliage is randomly distributed throughout the forest canopy (i.e. the instrument is based on Beers Law). The third and fourth assumptions are that the foliage elements are small and that foliage is randomly oriented (LI-COR 1991). Forests and plantation canopies do not conform to these assumptions because tree distribution, branches, and leaves are usually clumped. In addition leaves for some species are semi-transparent, and leaves and branches can partially reflect some light into the sensor leading to the underestimation of LAI. Consequently, LAI will be seriously underestimated by the LI-COR LAI2000 if it is not calibrated for a given species and stand structure to transform measurements of PAI to LAI.

Estimates of plant area as measured with a canopy analyser can be calibrated to estimates of LAI using a correction factor that takes into account non-photosynthetic mass such as branches and twigs. Cherry *et al.* (1998) calibrated estimates of LAI for eucalypt plantations using a LI-COR LAI-2000 and found LAI = 1.54*PAI -0.1. This relationship was established by comparing measurements of LAI with the LAI-2000 to destructively sampled trees.

Other methods of calibrating the LAI-2000 have been proposed by Gower and Norman (1991) who suggested multiplying the observed LAI value of each conifer species in a mixed stand by the corresponding total-projected-needle-area-to shoot-silhouette-area-ratio. This method is endorsed by the makers of the LAI-2000 in the instruments instruction manual (LI-COR-1991), not surprisingly as John Norman was the original developer of the instrument. The shoot ratio is a measurement of clumping on the shoot and is defined as the projected one-sided area of all the needles belonging to a shoot divided by the silhouette area of that shoot with its needles still on. The technique of measuring the shoot ratio is described in Gower and Norman (1991). They observed a highly significant relationship ($r^2 = 0.96$, $P<.005$, with a total of 5 observations). The reported n = 5 in Gower and Norman (1991) may be inadequate to satisfy statistical assumptions of normality and may mean that their reported r^2 is not accurate. There may also be a bias in this study depending on what silhouette a canopy analyser sees (this may change with the direction of the sun) through its lens and the silhouette used to create the shoot ratio. Gower and Norman report ratios of the total projected needle area to the shoot-silhouette as (mean $\pm$ 1 SD) 1.5 $\pm$ 0.41 for red pine, 1.67 $\pm$ 0.35 for white pine, 1.49 $\pm$ 0.28 for European larch, and 1.60 $\pm$ 0.14 for Norway spruce. These published values for pines might have been used to calibrate LAI measurements in this study since they might have similar shape characteristics to radiata pine.

While the method proposed by Gower and Norman (1991) is promising, a later study by Deblonde *et al.* (1994), showed that the technique was not applicable in all stands and suggest factors that may limit the general applicability. Deblonde *et al.* (1994) found that a further correction factor should be applied for forest stands with a significant woody biomass component. One may not simply take the shoot ratios measured by Gower and Norman and multiply them by LAI-2000 measurements to obtain an absolute value of LAI, as was done by Grong *et al.* (1992).

2.2.2 Pipe Theory Model (predicting LAI from sapwood area)

LAI can also be estimated using previously defined relationships between leaf area and the area of sapwood at the base of the live tree crown (Waring *et al.,* 1982). The idea proposed by Waring *et al.* (1982) is based on a strong relationship between leaf area index and the sapwood area at the base of the live crown of a tree. Sapwood area can be

measured with an increment corer at breast height if the taper of the tree up to the live crown is accounted for. This method requires pre-determined allometric relationships between sapwood area and LAI for a given species, site, or silvicultural treatment. A relationship for radiata pine may be available in the form of a biochemically based model written by Arneth *et al.* (1999),which estimated net primary productivity of radiata pine from stem diameter growth measurements.

The pipe model theory of estimating leaf area was not used here because it would not add any information to empirical growth models used in the study. Since growth is already predicted as some function of stem diameter and basal area then adding terms of leaf area that were in turn estimated from stem characteristics would lead to autocorrelation of errors. It was intended in this study to have an estimate of stand LAI that was independent of standard plot measurements of tree diameter, tree height, and stand basal area.

2.2.3 Measurement of Leaf area using a Sunfleck Ceptometer

LAI can be estimated indirectly by measuring photosynthetically active radiation using a portable integrating radiometer (Sunfleck Ceptometer, Decagon Devices Pullman WA, USA, 1987). A ceptometer measures photosynthetically active radiation in the range of 400-700 nm with 80 sensors and automatically takes an arithmetic average of these 80 readings. LAI is computed using the Beer-Lambert Equation (Equation 1.0), which is dependent on published values of light extinction *k* for a given species.

Anderson (1966) noted that *k* is not constant unless the angle of leaves in the canopy with respect to the ground is 0, and should vary by plant and tree species as stated previously. The light extinction coefficient *k* also varies with canopy structure and even with time of day under conditions of direct as opposed to diffuse radiation. Some values for light extinction have been published (Jarvis and Leverenz, 1983; Pierce and Running, 1988).

2.2.4 Measurement of Leaf Area using Fisheye Photography

Hemispherical canopy photography is one indirect optical technique that has been widely used in studies of canopy structure and forest light transmission. Photographs taken skyward from the forest floor with a 180° hemispherical (fisheye) lens produce circular images that record the size, shape, and location of gaps in the forest overstorey. Digital scanners or cameras convert these hemispherical images into bitmaps, which are then analysed using specialized image analysis software. Image processing involves the transformation of image pixel positions into angular coordinates, the division of pixel intensities into sky and non-sky classes, and the computation of sky-brightness distributions. These data are subsequently combined to produce estimates of growing-season light transmission, as well as other measures more directly related to canopy structure, such as openness, leaf area, and sunfleck frequency (Frazer *et al.*, 1999). Fisheye imaging is best done under diffuse sky conditions, where there is contrast between foliage elements and sky; otherwise, a sunlit leaf could be mistakenly identified as sky, for example and LAI may be underestimated (Wells, 1990). Fisheye photography images can be analysed to calculate LAI and canopy gap fraction on both film and digital images. Hale and Edwards (2002) compared the techniques of analysing both types of images and concluded that the two approaches produce comparable results. Research has shown hemispherical photography to be an efficient and reliable method of LAI measurement in forest environments (Welles, 1990). Based on error analysis, Chen (1996) stated that optical methods, if combined with clumping analysis in coniferous stands hold the potential to provide LAI estimates that are more representative than direct estimates of LAI based on destructive sampling techniques. Jonckheere *et al.* (2004) have also suggested the use of a hemispherical photography with a digital image analysis as a method to overcome a number of technical problems related to indirect LAI estimation.

2.3 Mortality Models

Mortality is said to be anything but straightforward to model since forest ecosystems are so complex and it is nearly impossible to predict future climatic conditions (Woollons and Hayward, 1985). Mortality models in many growth and yield systems show poor predictive power because they do not include variables that describe stresses on tree vigor

and attempt to model mortality as a continuous function over time. Woollons and Hayward (1985) directly modelled mortality using a variant of a basic difference equation. They applied a prediction equation of total mortality to the initial stocking to create a Weibull distribution of diameter. Mason (1992) also tried to improve predictions of mortality by incorporating environmental variables. Woollons (1998) improved upon his earlier approach (Woollons and Hayward, 1985) by modelling mortality on even aged stands with a two-step process. He argued that a conflict between data models and mortality functions is inherent. The reason for this conflict is that many of the mortality equations used predict continuous loss of stems (per unit area), when in fact many permanent sample plot data contain records of no mortality over several years. In this situation a conventional mortality model will always over-predict stem death. Woollons proposed a two step regression model that first estimates the probability of mortality occurring. Then a mortality equation was developed using only plots where mortality occurred over a given period. Estimates from the second model were then reduced by a factor, equivalent to the probability of death occurring, acquired from the logistic equation. I believe the two approaches proposed by Mason (1992), and Woollons (1998) can be combined to predict mortality as a two step regression model that incorporates some environmental variables, such as climate data or root zone water balance.

2.4 Water Balance Modelling

Incorporating a model of root zone water balance into a classical growth and yield model may increase accuracy of tree growth prediction, and possibly make the model more robust under changing climatic conditions. Water balance is a well established indicator of growing potential and is said to be a main constraint to tree development on dryland sites.

Reduction in the availability of water will decrease stem growth through restricting physiological processes such as leaf area development, photosynthesis, and stomatal conductance (Boomsma and Hunter, 1990). Cell enlargement (or leaf expansion) is a process that is especially sensitive to water stress. Atwell *et al.* (1999) describe the process of plant or tree height growth and later phases of leaf expansion as depending mainly on cell enlargement, which depends heavily on turgor pressure of water in the

cell. Further, stem growth by carbon assimilation has been found to be influenced by increases in soil water availability by (Dupouey *et al.,* 1993; Livingston and Spittlehouse, 1993).

Models of water balance have been included as components of highly parameterised process-based models, and hybrid models to predict growth of mature *Pinus radiata* on dryland sites (Walcroft *et al.,* 1997; Arneth *et al.,* 1999; Landsberg and Waring, 1997), but have yet to be explicitly included in any classical growth and yield model. An index for water stress (the water stress integral) was used by Meyers (1988) to account for almost all of the variation in basal area increment of a stand subjected to a range of irrigation and fertilization treatments. Given the importance of water availability in regulating productivity and basal area growth on dryland sites, the water balance model may provide a useful process-based approach that can be incorporated into a classical growth and yield model to predict tree growth.

Water balance is based on the law of conservation of mass: any change in the water content of a given soil volume during a specified period must equal the difference between the amount of water added to the soil volume and the amount of water withdrawn from it. In other words, the water content of the soil volume will increase when additional water from outside is added by infiltration or capillary rise, and decreases when water is withdrawn by evapotranspiration or deep drainage (Zhang et al. 2002). Root zone water balance has been modelled to describe juvenile growth in *Pinus radiata* plantations with weed competition in New Zealand by (Watt *et al.,* 2003) as:

$$W_i = W_{i-1} + P_i - E_{ti} - E_{twi} - E_{bi} - E_{bwi} - E_{gi} - F_i \qquad (1.2)$$

where P_i is rainfall, E_{ti} the transpiration from the tree canopy, E_{twi} the evaporation from the wet tree canopy, E_{bi} the transpiration from the broom canopy, E_{bwi} the evaporation from the wet broom canopy, E_{gi} the evaporation from the soil surface, and F_i drainage from the root zone (Watt *et al.,* 2003). Surface runoff was assumed to be insignificant.

Another approach used to describe root zone water balance is included in the physiological growth model 3-PG (Landsberg and Waring, 1997). This model is a single layer soil-water-balance model that operates on a monthly time step. Monthly rainfall

(plus irrigation) is balanced against monthly evapotranspiration computed using the Penman-Monteith equation (Landsberg and Gower, 1997, p. 76.) The general equation is a simplified version that used by Watt *et al.* (2003), and can be written as;

$$\theta_T = \theta_{T-1} + P - I - E - D \qquad (1.3)$$

Where: θ_T is the root zone water balance at time T (mm), θ_{T-1} is the root zone water balance at time in the previous month (mm), P is precipitation (mm), I is the canopy interception (mm), E is evapotranspiration (calculated using the Penman-Monteith equation,) and D is the drainage of water from the soil (mm).

The model is initialized with soil water content = maximum available water (Θ mm) in the rooting zone. This is dependent on the water holding characteristics of the soil and the rooting depth of the trees (see Landsberg and Gower, 1997). This model will be discussed in detail in chapter 6 of this thesis.

CHAPTER 3

VALIDATION OF EXISTING GROWTH AND YIELD MODEL CanSPBL

3.1 Introduction

The forest growth and yield model CanSPBL (Zhao, 1999) was written by Weizhong Zhao as part of his PhD thesis for the University of Canterbury in 1999. The model was written to predict growth of plantation-grown radiata pine on the plains and hills surrounding Canterbury on the Selwyn Plantation Board Ltd. estate.

In this study, the model was examined quantitatively by assessing model behaviour with a more recent data set. The procedure involved graphical displays and statistical tests. Potential correlation was detected by inspection of graphical plots of residual versus predictions and explanatory variables. Model residual errors or the difference between predictions made with CanSPBL and observations of growth, which are part of the current data set, will display certain trends along with initial conditions, projection length, or predicted values when the model is biased. This technique was also used by Zhao (1999) in his validation of an existing model CANTY.

Model performance has been described using average model bias (AMB) and model efficiency (EF) (Loague and Green, 1991; Zhao, 1999). Average model bias (Equation 3.0) is an average of errors for all predictions. An AMB of 0 would indicate a model with no bias. Model efficiency (Equation 3.1) is also a measure of model performance: a high value of EF (maximum value of 1) indicates a model of perfect fit, an EF value of 0 indicates a model of poor fit where the average value would model the relationship as well, and finally a negative EF result indicates an even poorer fit than the average value.

$$AMB = \frac{1}{n}\Sigma\left(Y_i - \hat{Y}_i\right) \quad (3.0)$$

$$EF = 1 - \frac{\Sigma\left(Y_i - \hat{Y}_i\right)^2}{\Sigma\left(Y_i - \overline{Y}_i\right)^2} \quad (3.1)$$

Where: Y_i is the observed, $\hat{Y}_i$ is the modelled, and $\overline{Y_i}$ is the average observed value.

Data used for validation of CanSPBL were very similar to those used for the original study by (Zhao, 1999), but were independent in that they included measurements for years and plots not in the previous analysis. Validation data were collected using measurement procedures that were the same as those used by Zhao (1999). This dataset covers forest types on plains, hills, and coastal sands with elevations ranging from 2 – 600 m above sea level. The forest types in these areas varied because of differences in both elevation and soil type. As forests are located further from the sea in the Canterbury Plains, soils change from coastal sands, shallow and dry floodplains soils, to deep wet LOESS hill soils (Barringer *et al.,* 1998). The validation exercise carried out using these data provided an opportunity to examine model bias in predicting future growth under possibly different environmental conditions, and may give some insight into the sources and magnitude of model bias. Verification in a qualitative sense will examine the biological aspects of each equation used for prediction and the structure of the model as a whole.

3.2 The existing model CanSPBL

The model CanSPBL is a non linear least squares regression system: the theoretical aspects of the system have been documented in Zhao (1999). The modelling approach used was to create both stand and tree level models. Stand level predictions made by CanSPBL (Zhao, 1999) were mean top height (MTH), basal area per hectare (G), stems per hectare (N), volume per hectare (V) and diameter distribution. Equations for each predicted variable were fitted using non-linear least-square regression procedures. Model performance was tested by examining graphical residual patterns. Diameter class distributions were described using a reverse Weibull function. Stand tables were

produced using a recovery method of parameters to project future stand statistics, and finally the method of moments was used to convert stand statistics of standard deviation, maximum value, and arithmetic mean diameter to Weibull distribution parameters. The following is a list of the equations used for both stand and tree level prediction and distribution recovery.

Table 3.00: Equations used for projection of stand tables and stand level predictions with CanSPBL

Number and name		Equation form
1.	Weibull density	$f(D)=\frac{c}{b}\left[\left(\frac{a-D}{b}\right)^{c-1} e^{\left[-\left(\frac{a-D}{b}\right)^{c}\right]}\right]$ (D = dbh class, a, b, c = parameters)
2.	Weibull-cumulative	$f(D)=e^{\left[-\left(\frac{a-D}{b}\right)^{c}\right]}$ (D = dbh class, a, b, c = parameters)
3.	Std. deviation of stand	$d_{std}=\sqrt{\frac{\left[\frac{1}{m\bar{n}}\sum\sum d_{pi}^2+\frac{1-\frac{m}{M}}{m\bar{n}(m-1)}n_p\bar{d}_p^2-\frac{1-\frac{1}{M}}{m\bar{n}^2(m-1)b}\left(\sum n_p\bar{d}_p\right)^2\right]}{1-\frac{1}{\bar{n}M}}}$ (d_{pi}=i^{th} dbh in p^{th} plot, m=sample size, M=potential plots, n_p=the number of trees in p^{th} plot, $\bar{n}$=average number of trees)
4.	Arithmetic mean	$\bar{d}=\sqrt{d_{\bar{g}}^2-d_{std}^2}=\sqrt{\frac{4000G}{\pi N}-d_{std}^2}$ ($d_{\bar{g}}$=diameter of mean basal area, G=basal area/ha, N=stems/ha)
5.	Maximum Dbh	$d_{\max}=\max(d_{pi})$ ($d_{pi}=i^{th}$ dbh in p^{th} plot)
6.	Location a	$a=d_{\max}$
7.	Scale b	$b=\left[\Gamma\frac{(1+1/c)}{(a-\bar{d})}\right]^{-1}$

8. Shape c

$$c = \left\{ z \left[1 + (1-z)^2 \begin{pmatrix} -0.2200991 - 0.00194664z + \\ 0.15310925z^2 - 0.08354348z^3 + \\ 0.007454537z^5 \end{pmatrix} \right] \right\}^{-1},$$

$$z = \frac{d_{std}}{\left(a - \overline{d}\right)}$$

9. Projection of Std. deviation of dbh

$$d_{std2} = e^{\left[\ln(d_{std1}) \left(\frac{T_1}{T_2} \right)^{\beta} + \frac{(\alpha_0 + \alpha_1 Alt)}{1000} \left(1 - \left(\frac{T_1}{T_2} \right)^{\beta} \right) \right]}$$

(T_1, T_2 = initial and projection age, Alt = altitude)

10. D_{max} projection

$$d_{max2} = e^{\left[\ln(d_{max1}) \left(\frac{T_1}{T_2} \right)^{\beta} + \frac{(\alpha_0 + \alpha_1 Alt)}{1000} \left(1 - \left(\frac{T_1}{T_2} \right)^{\beta} \right) \right]}$$

11. Height curve

$$h = 1.40 + \left[\frac{0.695955 + 0.666983T^{-0.5} - 0.106771\ln(SI) + }{\frac{(0.954201 + 0.000741Alt)}{d}} \right]^{-5}$$

(h = height, T = age, SI = site index, Alt = altitude, d = diameter)

12. Mean Top Height

$$H_2 = e^{\left[\ln(H_1) \left(\frac{(T_1 + t_0)}{(T_2 + t_0)} \right)^{\beta} + \frac{(\alpha_0 + \alpha_1 Alt + \alpha_2 (Alt - 250)X)}{1000} \left(1 - \left(\frac{(T_1 - t_0)}{(T_2 - t_0)} \right)^{\beta} \right) \right]}$$

(H_1= initial height, T_1, T_2 = initial and projection age, Alt = altitude , α_0 , α_1 , α_2 , t_0 , and β = parameters, X = dummy variable for altitude ($X = 0$ when $Alt < 250$, $X = 1$ when $Alt \geq 250$)

13. Basal Area

$$G_2 = e^{\left[\ln(G_1) \left(\frac{T_1}{T_2} \right)^{\beta} + \frac{(\alpha_0 + \alpha_1 Alt + \alpha_2 (Alt - 250)X)}{1000} \left(1 - \left(\frac{T_1}{T_2} \right)^{\beta} \right) \right]}$$

(G_1= initial basal area, T_1, T_2 = initial and projection age, Alt = altitude , α_0 , α_1 , α_2 , and β = parameters, X = dummy variable for altitude ($X = 0$ when $Alt < 250$, $X = 1$ when $Alt \geq 250$)

14. Stems per hectare

$$N_2 = \left(N_1 + \frac{a}{100000} \left(T_2^b - T_1^b \right) \right)^{\left(\frac{1}{c} \right)}$$

(N_1 = initial stems per hectare, T_1, T_2 = initial and projection age, a, b, c = parameters)

15. Volume per hectare

$$V = \alpha G^{\beta} H^{\gamma}$$

		(G = basal area per hectare, H = mean top height per hectare, α, β, γ = parameters)

Table 3.10: Equations used for projection of tree level projections and disaggregative adjustments with CanSPBL

Number and name		Equation form
1.	Probability of stem death	$P(death) = \left(e^{\frac{1}{(\alpha_0 - \alpha_1 Rd_1 + \alpha_2 Alt + \alpha_3 N_1 (T_2 - T_1) - \alpha_4 SI)}} + 1 \right)^{-1}$ (d_i = initial diameter of the i^{th} tree in a plot, $\bar{d}$ = arithmetic mean diameter in the plot, T_1, T_2 = initial and projection age, SI = site index, α_0, α_1, α_2, and α_4 = parameters, R_d = relative diameter if the i^{th} tree in a plot = $\frac{d_i}{\bar{d}}$)
2.	Stems per hectare	$n_{2i} = n_{1i}[1 - p(death)]$ (n_{1i} = initial number of stems per hectare in a plot, $p(death)$ = probability of stem death)
3.	Projected individual stem diameter	$d_2 = \sqrt{\frac{1}{0.00007854} \frac{G_2}{N_2} \left[\frac{0.00007854 d_1^2}{\left(\frac{G_1}{N_1} \right)} \right]^{\left(\frac{T_2}{T_1} \right)^{\beta}}}$
4.	Projected individual stem height	$h = 1.40 + \left[\frac{0.695955 + 0.666983 T^{-0.5} - 0.106771 \ln(SI) + (0.954201 + 0.000741 Alt)}{d} \right]^{-5}$ (h = height, T = age, SI = site index, Alt = altitude, d = diameter)
5.	Projected individual tree volume	$v = e^{\left\{ \alpha \ln(d) + \beta \ln\left[\frac{h^2}{(h - 1.4)} \right] + \gamma \right\}}$ (h = height, d = diameter, α, β, γ = parameters)

6. Adjust projections of individual trees

adjust n_{2i} so that $\Sigma n'_{2i} = N_2$ $\quad n'_{2i} = n_{2i} + \frac{(N_2 - \Sigma n_{2i}) n_{2i} d_{2i}^{-1}}{\Sigma n_{2i} d_{2i}^{-1}}$

adjust d_{2i} so that $\Sigma g'_{2i} n'_{2i} = G_2$ $\quad d'_{2i} = d_{2i} \sqrt{\frac{G_2}{0.00007854 \Sigma d_{2i}^2 n'_{2i}}}$

adjust v_{2i} so that $\Sigma v'_{2i} n'_{2i} = V_2$ $\quad v'_{2i} = \frac{V_2}{\Sigma v_{2i} n'_{2i}} v_{2i}$

3.3 Methods

The model CanSPBL was applied to independent data in this study using software developed by Dr. Weizhong Zhao as Basic code within a Microsoft Excel spreadsheet. Version (1.0) was used. Model examination in this study consisted of a locally weighted least squares (LOESS) analysis of residual plots against explanatory variables (Cleveland, 1979; Cleveland and Devlin, 1988). Graphical plotting of residuals against all predictor variables was also included, and measures of model efficiency EF (efficiency factor) and AMB (average model bias) were calculated (Loague and Green, 1991).

To test for bias in CanSPBL an independent validation dataset was used to examine the difference between expected and observed values. Validation data were collected on the same forest estate, using a similar measurement procedure as modelled by Zhao (1999).Validation data comprised 4486 plot measurements, from ages 7 – 30 years (Figure 3.1 (a)), and elevations ranging from 5 – 685(m). Measurement intervals were chosen to describe all possible re-measurement intervals (Figure 3.1 (b)). Zhao (1999) showed that all possible measurement intervals gave the best fit with growth and yield data in Canterbury when short, long, and all possible intervals were considered.

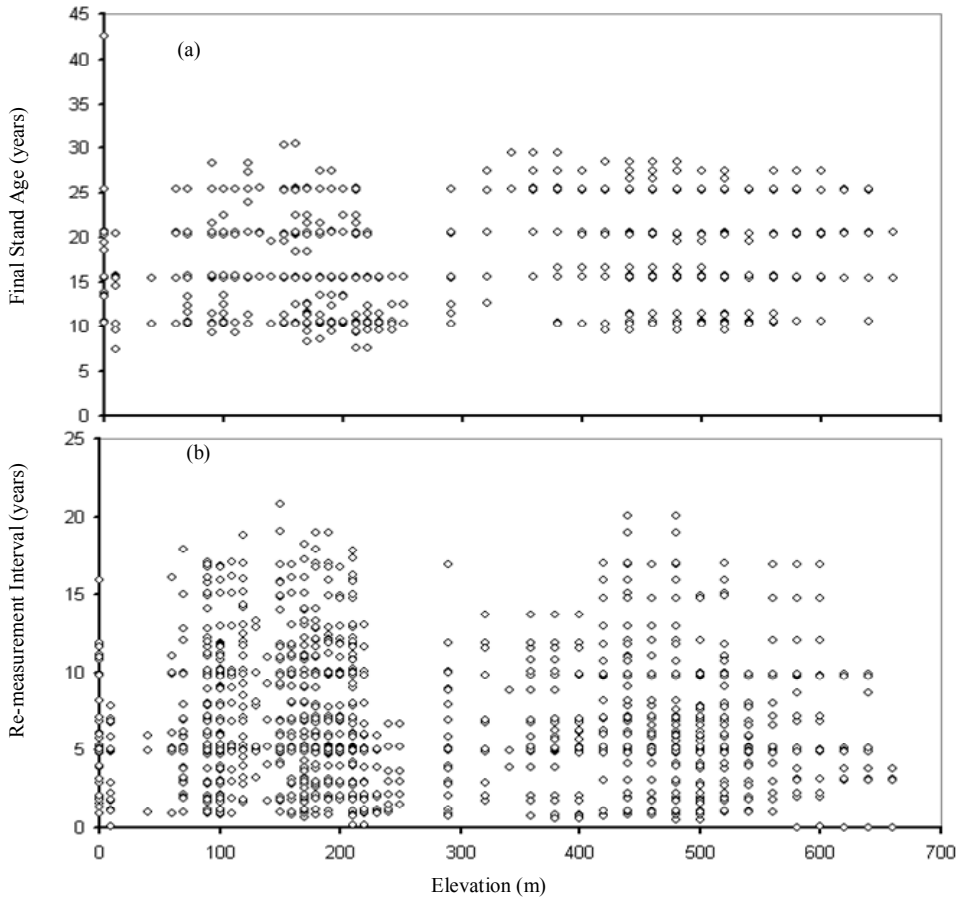


Figure 3.1: Validation data state space in terms of final stand age verses elevation (a), and re-measurement interval verses elevation (b).

The tested variables were the main components of the stand level model in CanSPBL; and were mean top height (m), basal area (sq m / ha), and stocking (stems / ha). The independent variables used for prediction were initial height, initial and projection age, initial basal area, and altitude. A binary variable of (0,1) representing sites above and below 250 m respectively was also used to implant a gradation with elevation.

3.4 Results

Predicted values were calculated with the model CanSPBL. Residuals are the difference between measured and modelled values. Graphical residual plots and frequency distributions are shown in Figures 3.2 – 3.10. For each modelled variable, five residual plots are shown; residuals versus predicted value, elevation, initial value, and initial and final stand ages. Average model bias and efficiency factors were calculated for each

model component and are listed in Table 3.1. A weighted least squares (LOESS) analysis was also used to look at the trend of residuals for each model component against elevation are shown in Figures 3.11 – 3.13.

LOESS stands for "locally weighted least squares." LOESS is a data analysis technique for producing a "smooth" set of values from a time series which has been contaminated with noise, or from a scatter plot with a "noisy" relationship between the 2 variables. This technique can provide insight into trends in scatter plots where data points are concentrated in regions of the x axis. In a time series context, the technique is an improvement over least squares smoothing when the data points are not equally spaced, which is assumed by least squares (Cleveland, 1979), and (Cleveland and Devlin, 1988). For LOESS smoothing, the analyst can vary the size of the smoothing window, given as the fraction (0 to 1). A smoothing of window 0.1 states that the window has a total width of 10% of the horizontal axis variable. A LOESS fit line in this analysis is intended to reveal trends in the residual plots of each model component with increasing elevation.

Table 3.1: Model performance measures for predicted mean top height (MTH), Basal area (G), and Stocking (N).

Model Component	Overall		Sands		Plains		Hills	
	AMB	EF	AMB	EF	AMB	EF	AMB	EF
MTH	0.460	0.942	1.744	0.898	0.418	0.948	0.442	0.940
G	-1.153	0.935	-2.177	0.728	-1.013	0.962	-1.499	0.944
N	-15.095	0.886	-125.822	0.096	-6.984	0.910	-28.484	0.924

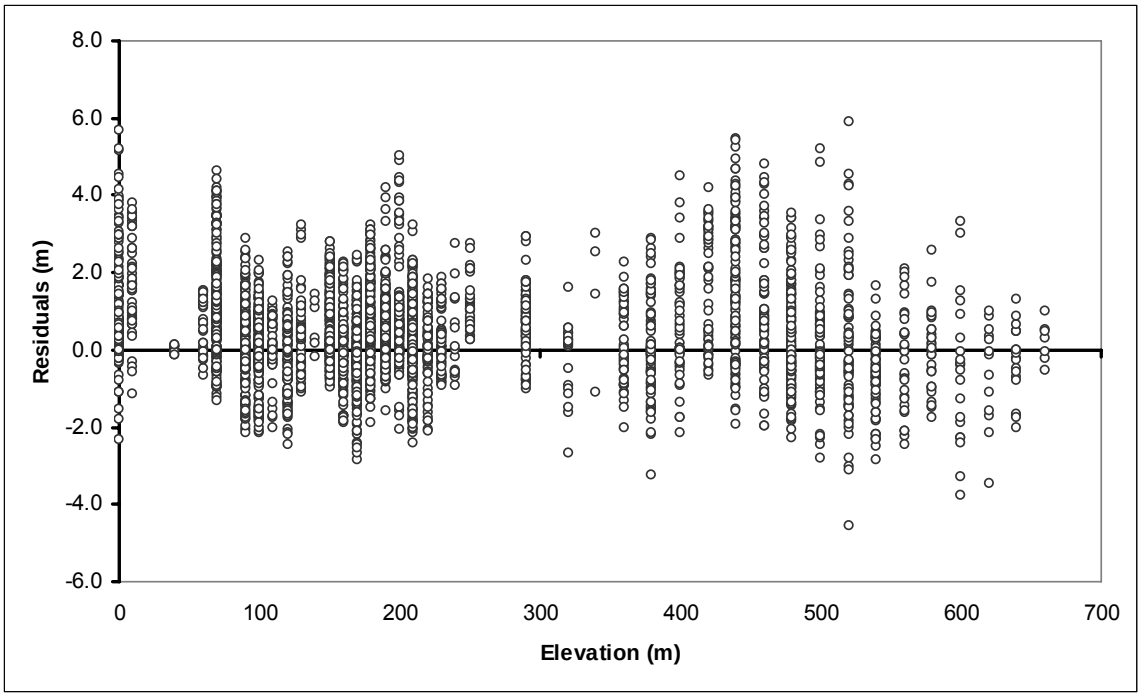


Figure 3.2: Residuals of projected mean top height verses elevation.

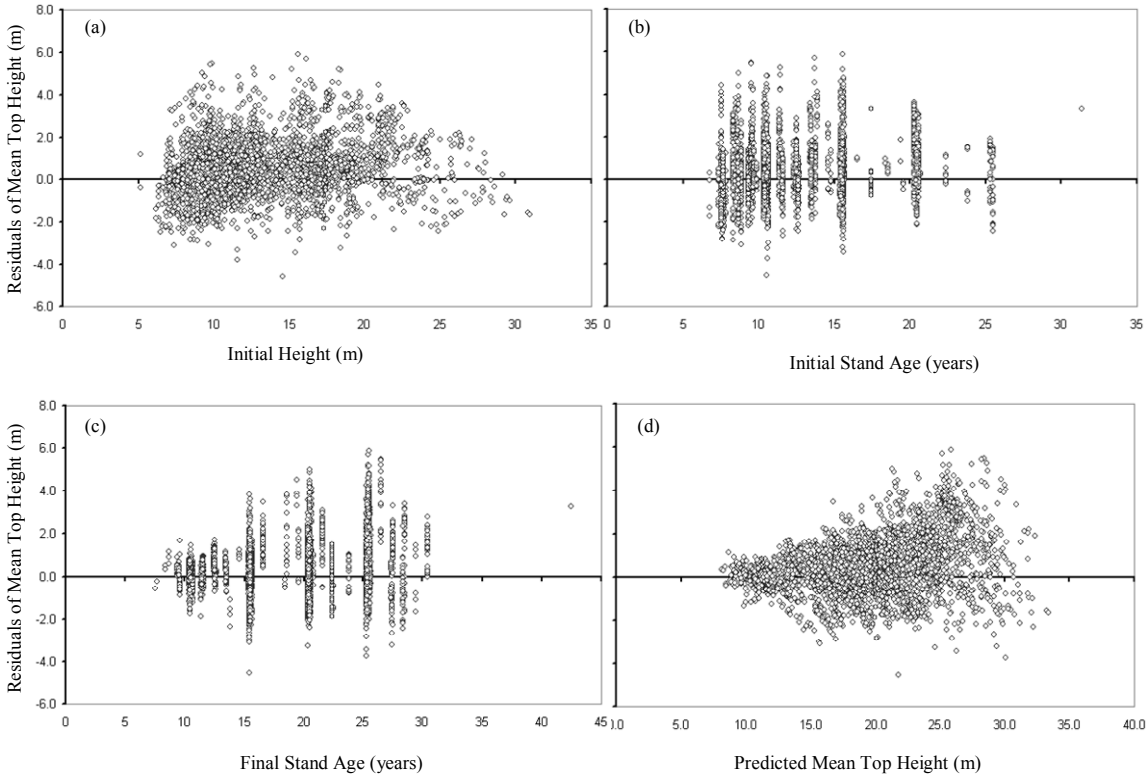


Figure 3.3: Residuals of projected mean top height in terms of residuals verses initial height (a), initial stand age (b), final stand age (c), and predicted mean top height (d)

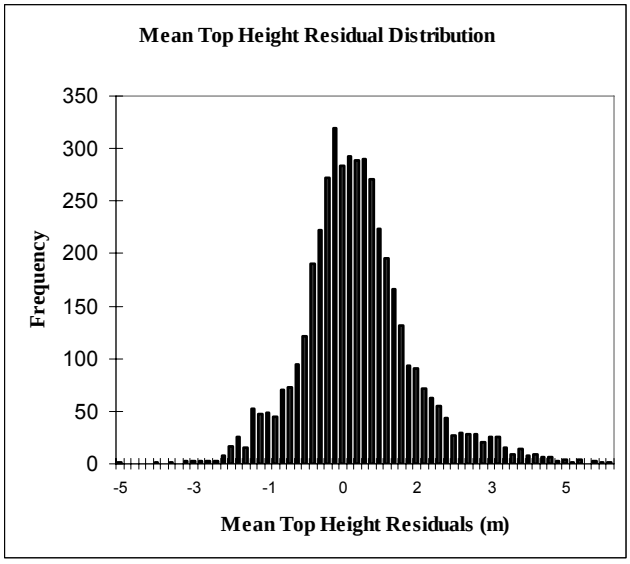


Figure 3.4: Mean top height residual frequency distribution.

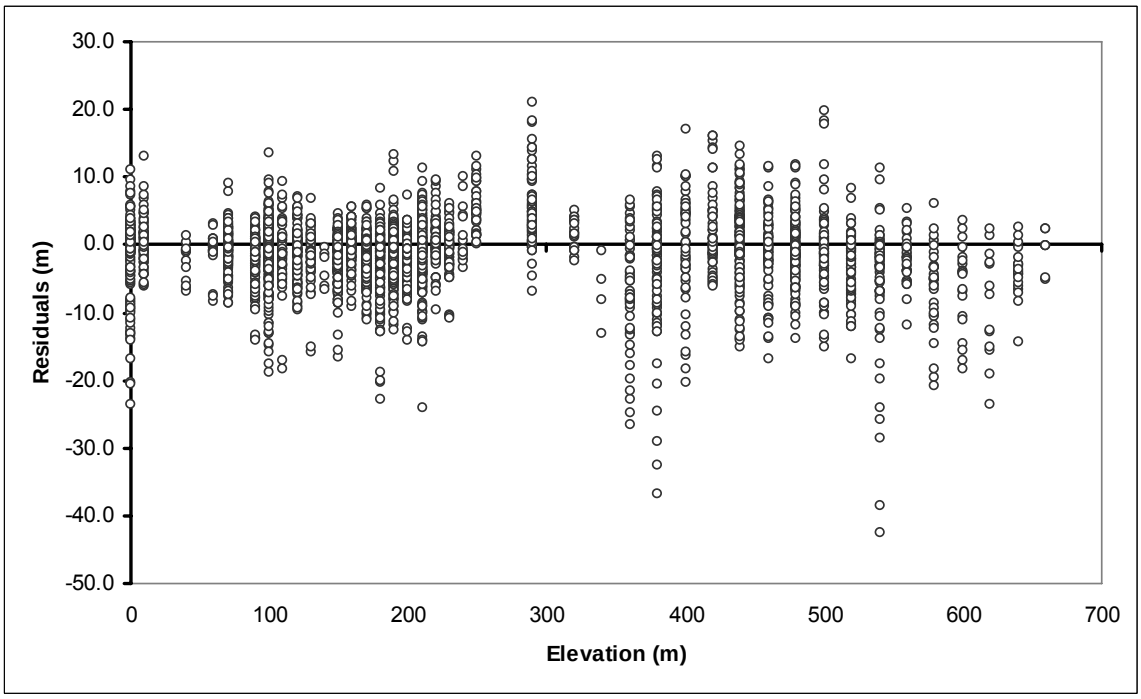


Figure 3.5: Residuals of projected stand basal area against elevation.

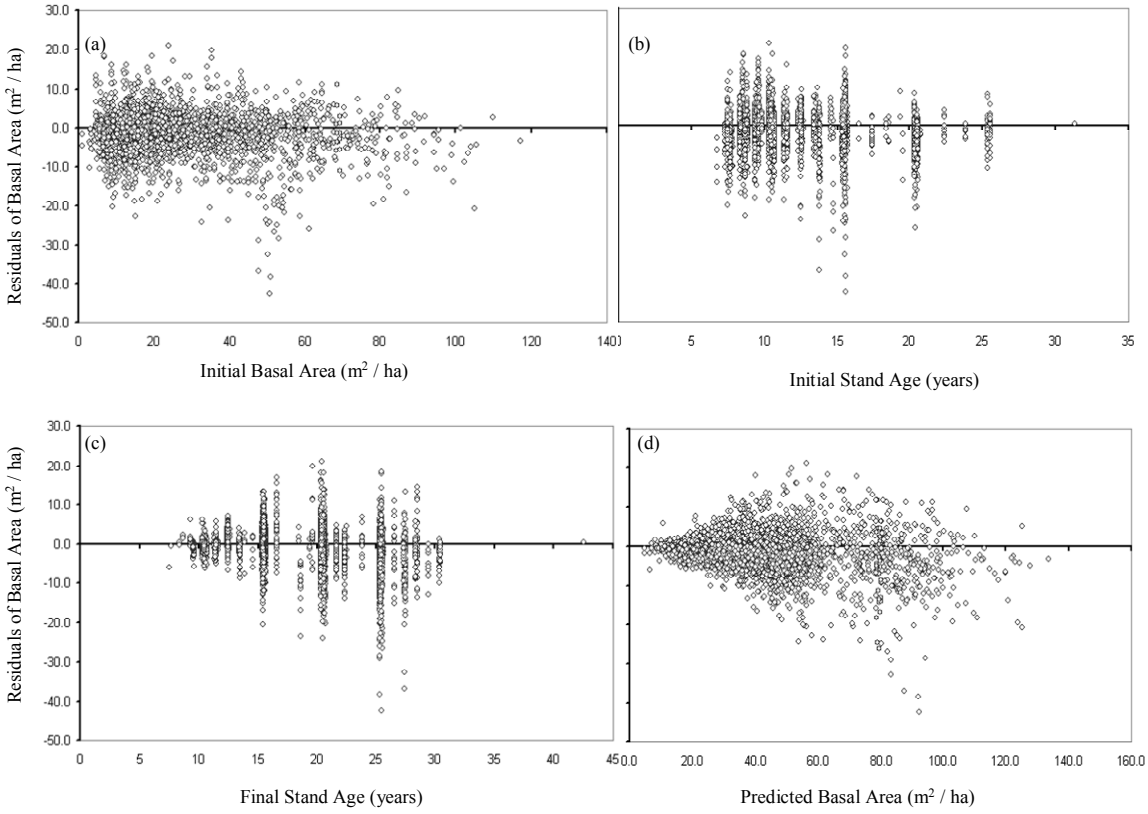


Figure 3.6: Residuals of stand basal area in terms of residuals verses initial basal area (a), initial stand age (b), final stand age (c), and predicted basal area (d).

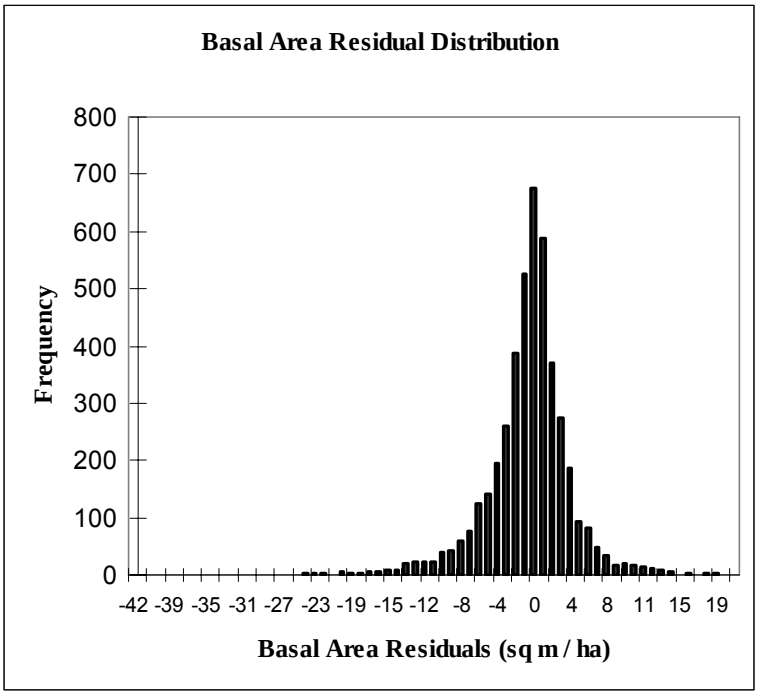


Figure 3.7: Basal area residual frequency distribution.

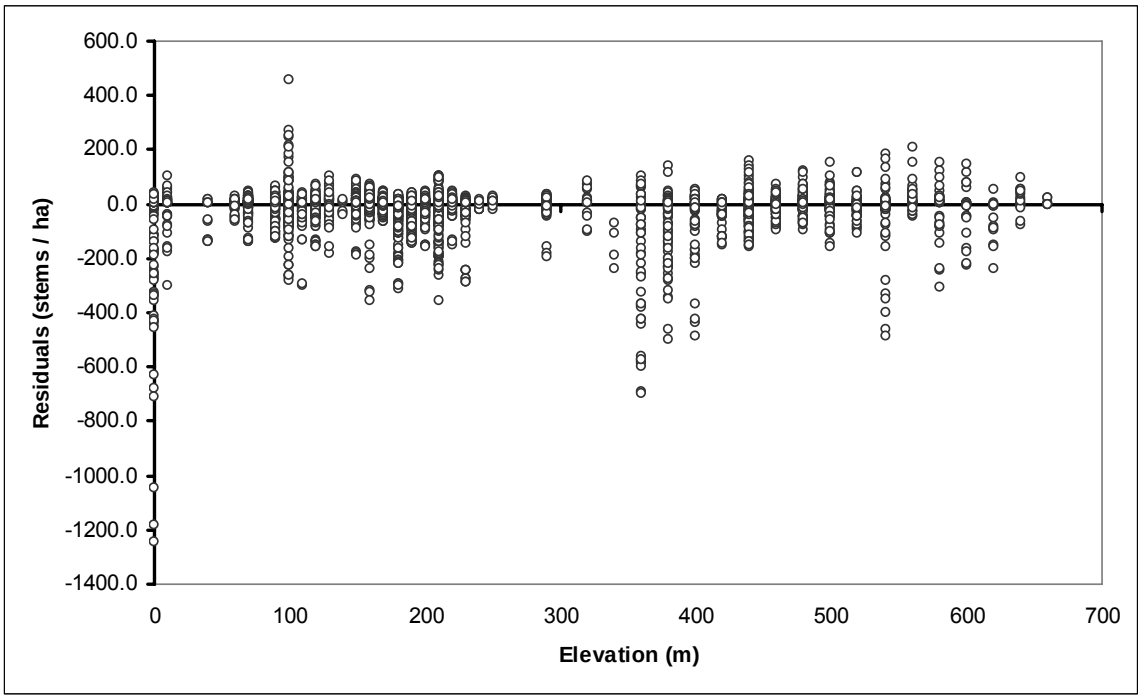


Figure 3.8: Residuals of projected stocking against elevation.

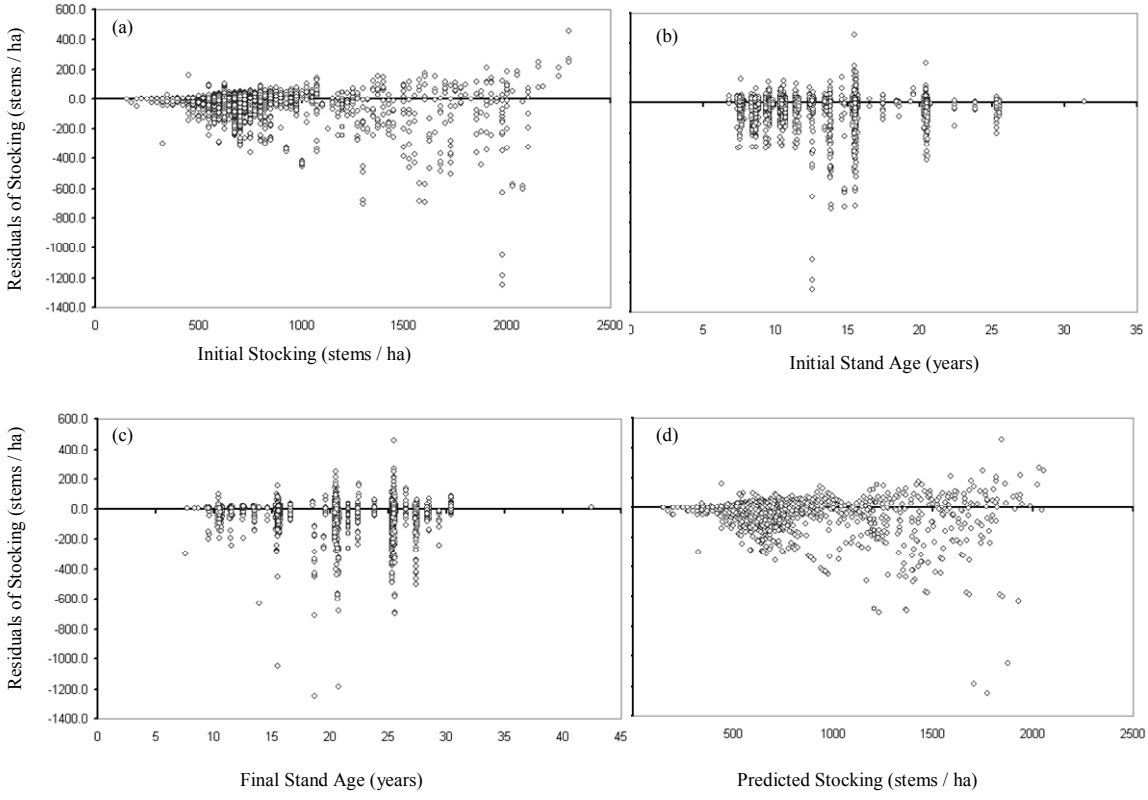


Figure 3.9: Residuals of stand stocking in terms of residuals verses initial stocking (a), initial stand age (b), final stand age (c), and predicted stocking (d)

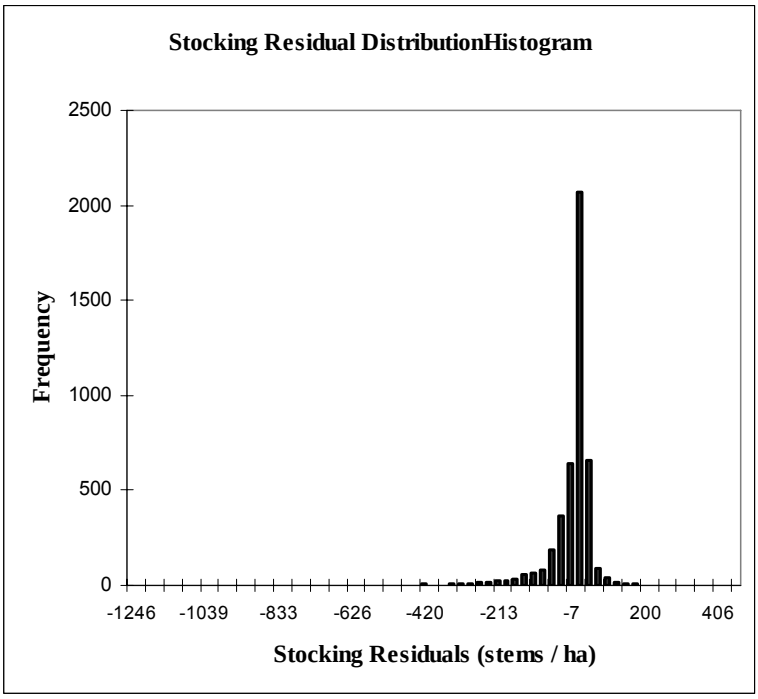


Figure 3.10: Stocking residual frequency distribution.

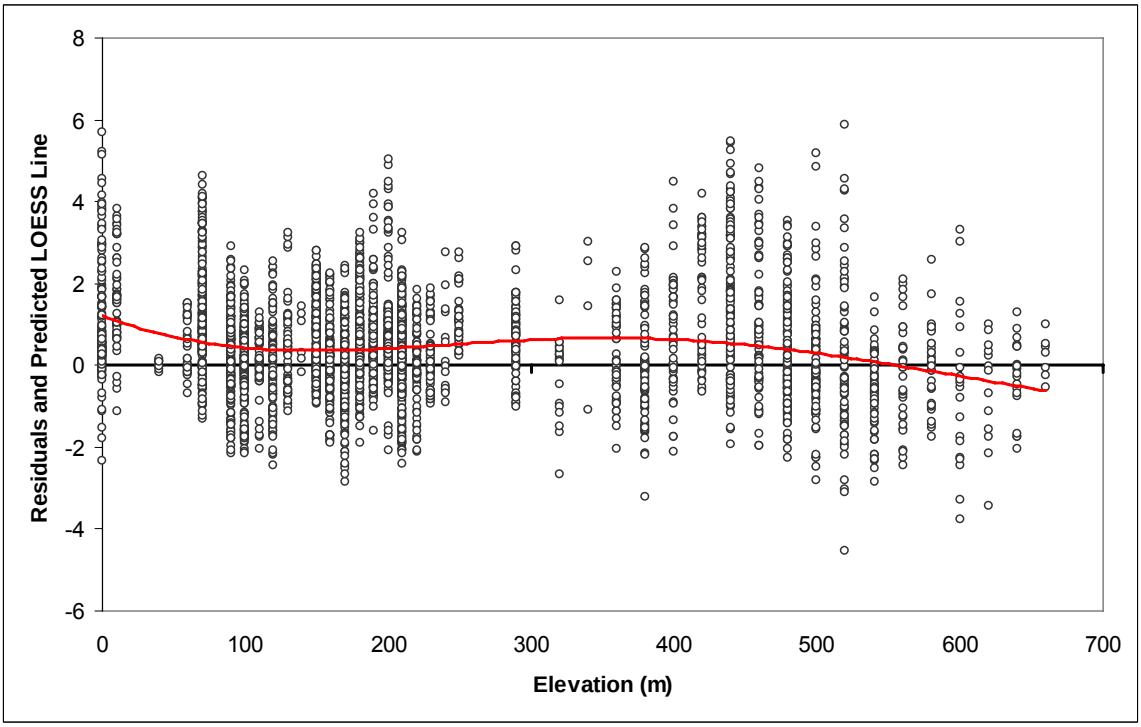


Figure 3.11: Residuals of projected mean top height (m) against elevation (m) plotted with LOESS regression line, f = 0.4.

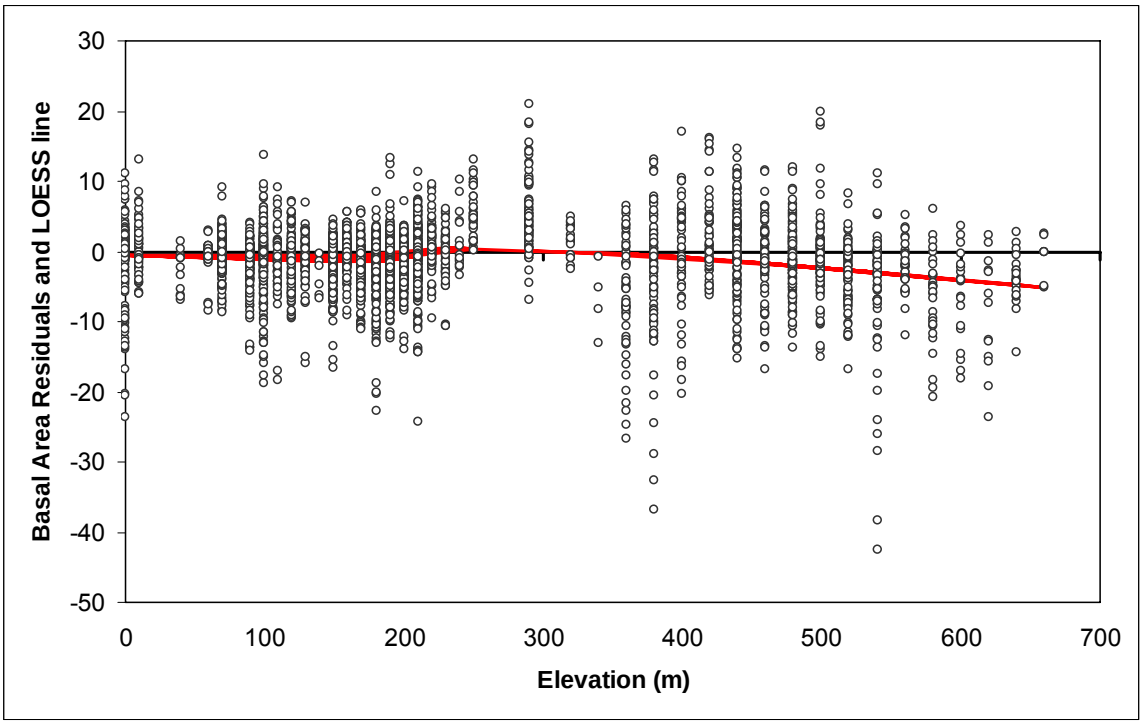


Figure 3.12: Residuals of projected Basal Area (sq m / ha) against elevation (m) plotted with LOESS regression line, f = 0.6.

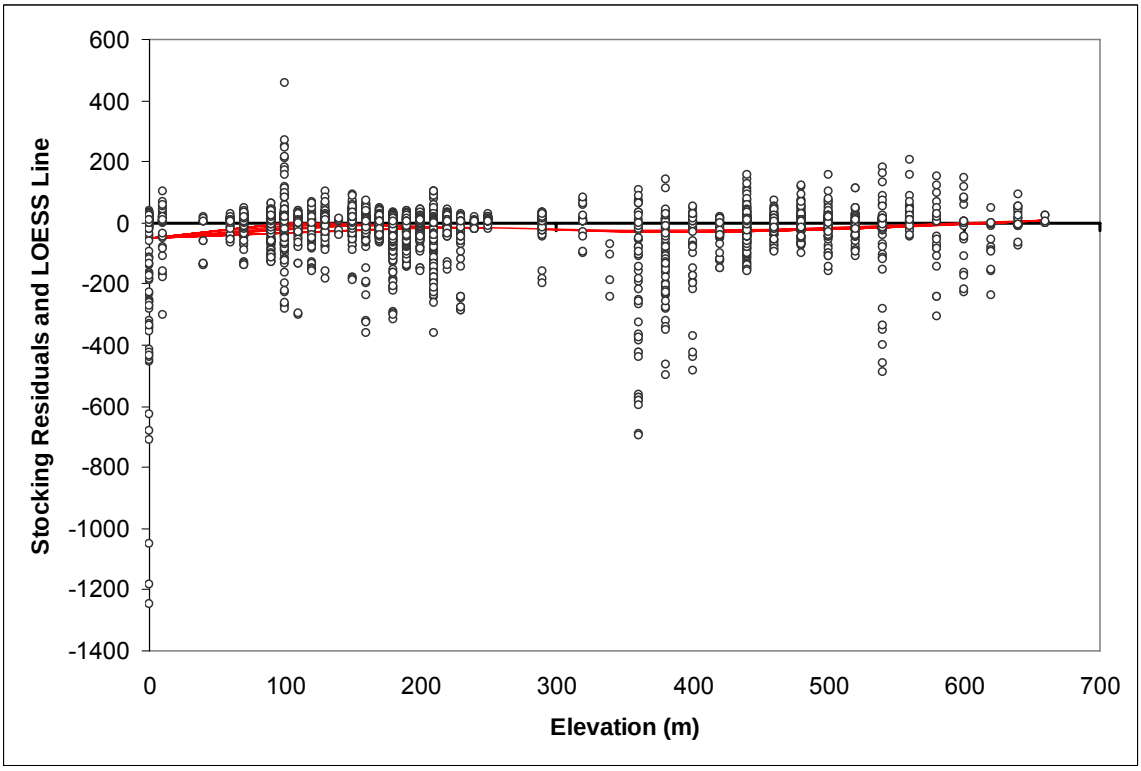


Figure 3.13: Residuals of projected Stocking (stems / ha) against elevation (m) plotted with LOESS regression line, f=0.6.

3.4 Conclusions

Examining residual plots for mean top height, basal area, and stocking showed bias in projection estimates that suggested there would be benefits in updating the existing model with a more current dataset that more completely covers the range in altitudes on the SPBL estate (Figures 3.2, 3.3, 3.5, 3.6, 3.8, and 3.9). Plots and model efficiency measures showed bias in under-prediction of mortality, under prediction of basal area, and over-prediction of mean top height. Residual plots for stocking (Figures 3.8 and 3.9) showed bias in under prediction of mortality, especially for lower and higher altitude forests. This bias in stocking prediction with respect to elevation was also shown with an (AMB) of -125.8 for sands, - 6.9 for plains, and -28.5 for hills plots. These same trends followed for model efficiency calculations at the three site types (Table 3.1). LOESS analysis on stocking residuals showed an increase in residuals (also indicating an under prediction in mortality) at elevations below 80m and above 250 (m) (Figure 3.13). The basal area model showed similar trends of bias in under prediction as the stocking model (Figures 3.5 and 3.6. Areas of bias were also concentrated in plots classified as sands or hills (Table 3.1). LOESS analysis on basal area showed bias at similar elevations but at a smaller magnitude than indicated by AMB and EF measures (Figure 3.12). However, residual plots for mean top height (Figures 3.2 and 3.3) showed a bias in over prediction of height, especially for lower and higher altitude forests. This trend in mean top height was also described by model efficiency measures increasing for sands and hills forest types (Table 3.1)

3.5 Discussion

This section covered a validation of CanSPBL (Zhao, 1999) using a new dataset. This new dataset included sites within the modelled forest estate that were not considered in the original model, these sites were generally higher elevation forests older than 15 years of age. The original model was not intended to distinguish between the three distinct forest types of sands, plains and hills. This lack of distinction highlights areas where bias was most likely to occur in stand level projections. An updated model is intended to encompass a larger state space and vary predictions of growth and stand volume according forest type.

CHAPTER 4

UPDATING CanSPBL

4.1 Introduction

The stand-level growth and yield model CanSPBL (Zhao, 1999) was shown in chapter 3 to have bias associated with mean top height, basal area, and stocking. These biased projection estimates suggest there would be benefit in updating the existing model with a more current dataset that completely covers the range in altitudes on the SPBL estate. A new model, (CanSPBL(1.2)) was developed using a more complete dataset than that used by Zhao (1999). The components of the model CanSPBL(1.2) at a stand level comprise mean top height (MTH), basal area per hectare (G), stems per hectare (N), and diameter distribution. The newly established model was validated and compared to CanSPBL (1.0) with the same validation dataset. The data source for validation was prepared at a plot level to test projections of MTH, basal area per hectare, stems per hectare, maximum diameter, and the standard deviation of diameter.

Mortality models are consistently the weak link in many statistically based forest growth models. The factors that cause death in trees in the Canterbury region include wind throw, insect and disease attacks, animals, and competition from neighbouring trees. These factors have historically been hard to predict precisely, thus the amount of surviving trees per hectare is a model that has been hard to accurately fit (Woollons, 1998: Zhao, 1999). There has been considerable effort to improve prediction of mortality in New Zealand by incorporating environmental variables (Mason, 1992) and using variants to basic difference equations (Woollons and Hayward, 1985).

In this study, many different approaches were tested to try and improve the accuracy of the mortality model. These included a two step logistic approach similar to that used by Woollons (1998), including the environmental variable maximum wind speed over the observation period as a predictor variable, and using a mortality severity index as a means to reduce the modelling dataset to only include non – catastrophic mortality events.

4.2 Methods and Procedures

General Methods

To fit equations, methods used by Zhao (1999), of non linear least-square procedures with SAS software (SAS Institute Inc., 2001) were employed. The Mean Square Error (MSE) and graphical residual patterns were used as the selection criteria to judge model performance. Skewness and kurtosis were checked for final models to determine the magnitude of residual distributions for normality. Plots of residuals versus predictions and all possible explanatory variables were inspected to check for trends but only the four most important graphs are displayed in this study for each selected equation. The four graphs are residuals versus prediction, age, elevation, and time increment. The effects of elevation have been revealed by Zhao (1999), and during tests of CanSPBL(1.0) in chapter 3, and were incorporated into modelling equations where appropriate. Plots of residuals versus elevation were the final graphs to show the lack of bias.

No statistical tests are given in this analysis because repeated measurements have been taken from the basic experimental units (PSP). The consequences of this are:(i) estimators of the regression coefficients may no longer have minimum variance but will still be unbiased and consistent; (ii) standard errors of coefficients in the regression will be underestimated; and, (iii) any significance tests or confidence limits constructed using t or F distributions are likely to be wrong since assumed independence of errors is violated (West *et al*., 1984). The MSE for the regression is also likely to be underestimated if the correlation is positive and inflated if the correlations are negative (Snowdon *et al.,* 1999). The approach of using auto correlated datasets for fitting sigmoid equations has been defended by Clutter *et al.*(1983), by pointing out that growth modellers are not tying to prove that forest growth is described by sigmoid curves. The use of sigmoid curves is taken as an a-priori assumption. Rather, modellers are only trying to condition these equations to get the best fit for a growth region.

Some statistically valid tests are provided in this analysis as a check of the final results by preparing an auto correlation free data set. An auto correlation free data set was used to test the significance of explanatory variables being different from zero for all models and check the normality of residuals with the Kolmogorov-Smirnov test. The Kolmogorov-Smirnov test was chosen to test normality because of the large size of the data set.

The auto correlation free data set was prepared in a way such that one pair of repeated measurements from each plot was randomly selected from each plot. Data with auto correlation due to repeated measurements create unbiased coefficients with least-squares regression, but underestimate both the variances of error terms and the variances of coefficients. Thus, hypothesis tests based on auto correlated data are invalid (Neter and Wasserman, 1974; West *et al.,* 1894; West, 1995), but there are no problems with parameter estimates obtained by fitting a model to such data.

A reverse Weibull function was used to describe diameter-class distributions at a stand level. To produce stand tables, the recovery method was used to project future from the current stand statistics and the method of moments was used to convert the stand statistics of standard deviation, maximum value, and arithmetic mean diameter to Weibull distribution parameters. The standard deviation of a stand was obtained from the cluster sampling method used by Garcia (1991). Maximum diameter was estimated from all plots in a stand. Projection equations for standard deviation and maximum diameter were derived. Arithmetic mean diameter was solved for from the outputs of projection equations of basal area and stocking.

To validate the newly established model and test its performance against the existing model CanSPBL(1.0) at a plot level, a validation data set was prepared. The main model components of MTH, basal area per hectare, stems per hectare, maximum diameter, and standard deviation of diameter were examined using data at a plot level. One source of data was selected of independent plot measurements chosen randomly before model fitting. The average model bias (AMB), efficiency factor (EF), skewness and kurtosis were calculated for each model component. Graphs of residual patterns were examined to detect bias. Normality of residual plots were also tested using the Kolmogorov-Smirnov test.

4.3 Analysis and Results

4.3.1 Description of the Data for Growth Modelling

A reliable model can be built only with data of the highest quality. A sample data set covering the population both spatially and temporally is needed for model construction (Vanclay, 1994). A model constructed with a database having inappropriate coverage may result in prediction bias.

The model CanSPBL(1.2) is a stand level model for SPBL's estate covering rotation ages of 7.5 to 30 years. Samples of modelling data have been collected covering a wide range of conditions. Following methods used by Zhao (1999), tables and graphs were used to display the quality of permanent sample plots for growth modelling. Three main graphs were used to display database characteristics for stands and plantations. The first was stocking versus tree size or age, as suggested by Vanclay *et al.* (1995) and used by Zhao (1999). The second two graphs used to examine the quality of modelling data were MTH and stand basal area versus stand age, as used by Zhao (1999) and suggested by Garcia (1984, 1988, 1994).

Data were screened to remove plots that were identified as having errors. The data set was also filtered to only include measurements where no thinning had occurred over the re-measurement interval. Plots were removed from the dataset that were found to have values that were not in agreement with those written on field datasheets. The database was randomly sampled and datasheets were inspected for consistency with database values. It was found that approximately 5% of all the plots in the data summaries contained some sort of data entry or difference calculation errors such as declining tree diameters and heights over time. The data set was also limited to stands that did not exceed 30 years of age, as 30 years was the standard rotation length. Stands over 30 years old were usually unmanaged stands on the urban wildland interface that had an uncharacteristic management regime. An outlier, a 42 year old stand that was more than 2 standard deviations away from the mean of both basal area and mean top height was

removed from the analysis. The software used to implement equations fitted from this dataset will be constrained to stands less than 30 years of age.

The entire database comprised an average of 6 plots per stand. The interval of re-measurement averaged 6.5 years, and 96% of all the plots were 0.04 ha in size, while the remaining 4% were 0.02 ha. The database was partitioned for model building and model validation. In all there were 4416 plot measurements in the entire data set: 3667 were randomly chosen for model building, while 969 plot measurements were randomly set aside for model validation. Table 4.0 summarises the plot variables in the data set for model building and table 4.1 for model validation. Figure 4.0 displays the plot development pattern of the main components of mean top height (MTH), basal area, and stems per hectare with data for model building and Figure 4.1 with data for model validation.

Data for model building

There were 3667 plot measurements within the 749 plots selected for model building. A summary of plot variable estimates with a breakdown for plains and foothills plots are listed in Tables 4.0 and Figure 4.0 displays growth patterns of the main variables. Data for model building are summarised in Table 4.0.

Table 4.0: Summary of data for model building. Data shown represents 2888 plot measurements taken from plots located in the plains, and 779 plot measurements taken from the hills.

Variable: **Plains Data**	Units	Mean	Standard Deviation	Minimum	Maximum
Age (Initial)	Years	11.6	3.7	7.3	25.5
Age (Final)	Years	18.1	5.1	7.7	30.4
Stocking (Initial)	Stems / ha	679.5	180.5	200	2300
Stocking (Final)	Stems / ha	652.9	168	200	2300
Mean Top Height (Initial)	m	12.7	3.9	6.9	29.1
Mean Top Height (Final)	m	19.4	4.7	8.5	30.9
Basal Area (Initial)	m^2 / ha	19.7	10.6	3.4	66.6
Basal Area (Final)	m^2 / ha	35.3	12	5.7	69.2
Dbh Std. D (Initial)	cm	2.9	1.3	0.8	9.7

Dbh Std. D (Final)	cm	4.4	1.6	0.9	11.3
Dbh Max. (Initial)	cm	24	6.7	12.5	53.3
Dbh Max. (Final)	cm	33.8	7.1	15.8	57.1
Elevation	m	140.5	60.4	0	250

Variable: **Hills Data**	Units	Mean	Standard Deviation	Minimum	Maximum
Age (Initial)	Years	13	4.4	7.6	25.5
Age (Final)	Years	19	5.7	9.6	29.4
Stocking (Initial)	Stems / ha	871.3	438.5	150	2250
Stocking (Final)	Stems / ha	789.7	359.4	150	2175
Mean Top Height (Initial)	m	13.6	5	5.2	30.9
Mean Top Height (Final)	m	20.6	6.3	8.1	34.3
Basal Area (Initial)	m^2 / ha	35.6	23.2	1.1	117.2
Basal Area (Final)	m^2 / ha	57.3	24.6	3.1	130.5
Dbh Std. D (Initial)	cm	4.1	1.9	1.2	10.4
Dbh Std. D (Final)	cm	5.9	2.2	1.2	12.8
Dbh Max. (Initial)	cm	29.4	8	12.4	60.2
Dbh Max. (Final)	cm	41	8.5	17.6	65.4
Elevation	m	453	83.2	290	660

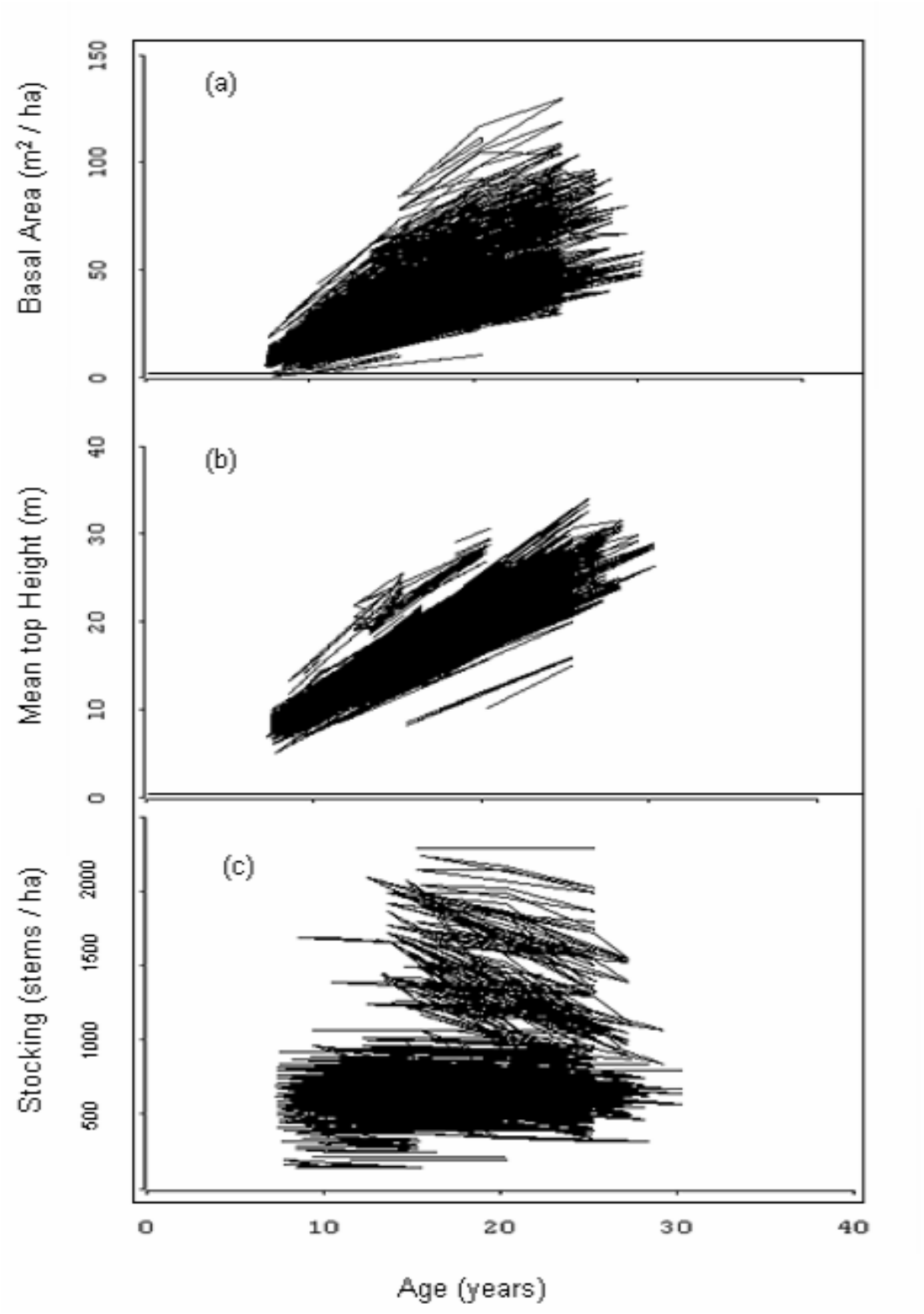


Figure 4.0: Growth patterns over time of (a) basal area, (b) mean top height, (c) stems per hectare using data for the whole model building data set.

Data for Model Validation

There were 969 plot measurements within the195 plots selected for model validation. A summary of plot variable variables with a breakdown for plains and foothills plots are listed in Table 4.1, and Figure 4.1 displays growth patterns of the main variables. Data for validation are summarised in Table 4.1.

Table 4.1: Summary of data for model validation. Data shown represents 728 plot measurements taken from plots located in the plains, and 241 plot measurements taken from the hills.

Variable: **Plains Data**	Units	Mean	Standard Deviation	Minimum	Maximum
Age (Initial)	Years	11.5	3.8	7.5	25.5
Age (Final)	Years	18	5.2	7.7	30.4
Stocking (Initial)	Stems / ha	668.1	143	275	1600
Stocking (Final)	Stems / ha	645	137.4	275	1600
Mean Top Height (Initial)	m	12.4	3.7	7.2	25.9
Mean Top Height (Final)	m	19.1	4.9	8.7	29.3
Basal Area (Initial)	m^2 / ha	19.1	11.4	4.6	61.6
Basal Area (Final)	m^2 / ha	35.1	13.6	7.2	73.1
Dbh Std. D (Initial)	cm	2.8	1.2	0.8	8.2
Dbh Std. D (Final)	cm	4.4	1.5	0.9	8.7
Dbh Max. (Initial)	cm	23.4	6.3	13	48.2
Dbh Max. (Final)	cm	33.4	7.2	15.7	50.5
Elevation	m	135.8	60.1	0	250

Variable: **Hills Data**	Units	Mean	Standard Deviation	Minimum	Maximum
Age (Initial)	Years	12.9	4.6	7.6	25.5
Age (Final)	Years	19.6	5.7	9.6	29.4
Stocking (Initial)	Stems / ha	823.2	424	250	2100
Stocking (Final)	Stems / ha	754.7	362.7	250	1975
Mean Top Height (Initial)	m	13	5.2	6.5	30.8
Mean Top Height (Final)	m	21.1	6.8	8.4	33.9
Basal Area (Initial)	m^2 / ha	34.7	26.9	4.9	104
Basal Area (Final)	m^2 / ha	59.7	27	9.2	123
Dbh Std. D (Initial)	cm	4.1	1.8	1.9	12.2
Dbh Std. D (Final)	cm	6.3	2.4	2.2	13.1
Dbh Max. (Initial)	cm	29.6	9.6	16.5	63.6
Dbh Max. (Final)	cm	43.4	10	17.6	69.3
Elevation	m	455.6	85.6	290	660

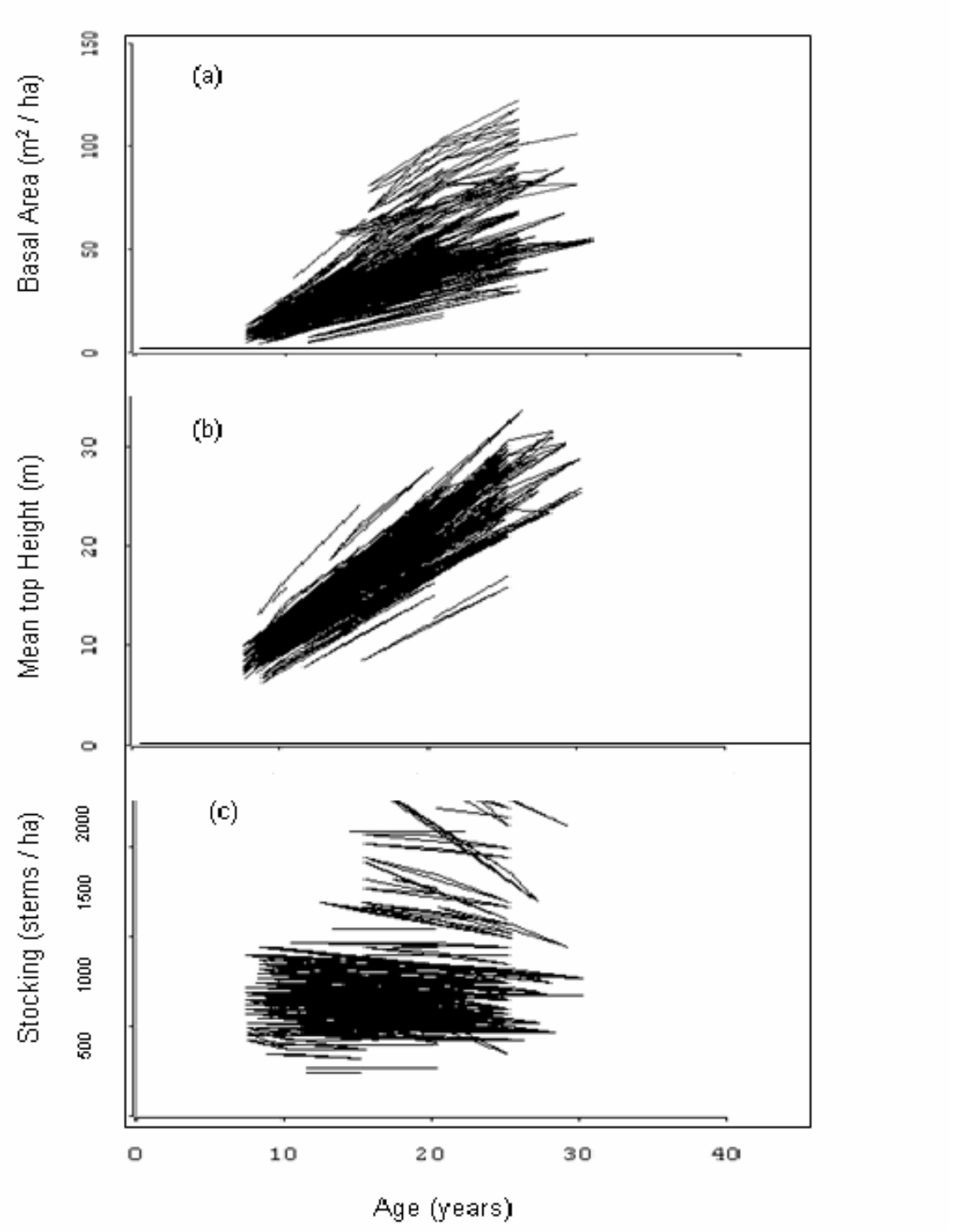


Figure 4.1: Growth patterns over time of (a) basal area, (b) mean top height, (c) stems per hectare using data for the whole model validation data set.

4.3.2: Modelling mean top height and basal area per hectare

All (sigmoid) equations in table 4.2 were fitted with mean top height and basal area data. The results of initial model fitting of mean top height and basal area with difference equations are listed in table 4.3 in terms of MSE. The equations with the 4 lowest mean square errors were tested with other predictor variables such as elevation and aspect to see if fit could be improved as was done in Zhao (1999). The final model chosen was the equation form having the lowest mean square error after effects of elevation were added. Final models were also chosen through the examination of residual plots against explanatory variables to look for possible bias. The distribution of residuals was examined using normal probability plots, and were tested for normality using the Kolmogorov-Smirnov test.

Table 4.2: Difference equation forms for mean top height and basal area.

Difference Equation Forms
Polymorphic Schumacher I: $Y_2 = e^{\ln(Y_1)\cdot\left(\frac{T_1}{T_2}\right)+\alpha\cdot\left(1-\left(\frac{T_1}{T_2}\right)\right)}$
Polymorphic Schumacher II: $Y_2 = e^{\ln(Y_1)\cdot\left(\frac{T_1}{T_2}\right)^{\beta}+\alpha\cdot\left(1-\left(\frac{T_1}{T_2}\right)^{\beta}\right)}$
Polymorphic Schumacher III: $Y_2 = Y_1^{\left(\frac{T_1}{T_2}\right)} e^{\alpha\left(1-\frac{T_1}{T_2}\right)^{\gamma}}$
Anamorphic Schumacher I: $Y_2 = Y_1 \cdot e^{\left(-\beta\cdot\left(\frac{1}{T_2^{\gamma}}-\frac{1}{T_1^{\gamma}}\right)\right)}$
Anamorphic Schumacher II: $Y_2 = Y_1 e^{-\beta\left(\left(\frac{1}{T_2}\right)^{\gamma}-\left(\frac{1}{T_1}\right)^{\gamma}\right)}$
Polymorphic Hossfeld: $Y_2 = \left[\left(\frac{1}{Y_1}\right)\cdot\left(\frac{T_1}{T_2}\right)^{\beta}+\alpha\cdot\left(1-\left(\frac{T_1}{T_2}\right)^{\beta}\right)\right]$
Anamorphic Hossfeld: $Y_2 = \frac{1}{\left((1/Y_1)+\alpha(1/T_2{}^{\beta}-1/T_1{}^{\beta})\right)}$

Anamorphic Weibull: $Y_2 = Y_1 \cdot \left[\frac{1-e^{-\beta \cdot T_2^{\gamma}}}{1-e^{-\beta \cdot T_1^{\gamma}}}\right]$

Polymorphic Weibull I: $Y_2 = Y_1 \cdot e^{\left(-\beta \cdot \left(T_2^{\gamma} - T_1^{\gamma}\right)\right)} + \alpha \cdot \left(1 - e^{\left(-\beta \cdot \left(T_2^{\gamma} - T_1^{\gamma}\right)\right)}\right)$

Polymorphic Weibull II: $Y_2 = \alpha - \beta\left(\frac{\alpha - Y_1}{\beta}\right)^{\left(\frac{T_2}{T_1}\right)^{\gamma}}$

Anamorphic Gompertz: $Y_2 = Y_1 \cdot \frac{e^{\left[-\beta \cdot e^{(-\gamma \cdot T_2)}\right]}}{e^{\left[-\beta \cdot e^{(-\gamma \cdot T_1)}\right]}}$

Polymorphic Gompertz I: $Y_2 = e^{\ln(Y_1) e^{-\beta (T_2 - T_1)} e^{\alpha (1 - \beta (T_2 - T_1))}}$

Polymorphic Gompertz II: $Y_2 = e^{\left[\ln(Y_1) \cdot e^{\left(-\beta \cdot (T_2 - T_1) + \gamma \cdot \left(T_2^2 - T_1^2\right)\right)}\right]} \cdot e^{\left[\alpha \cdot \left(1 - e^{\left(-\beta \cdot (T_2 - T_1) + \gamma \cdot \left(T_2^2 - T_1^2\right)\right)}\right)\right]}$

Polymorphic Von Bertalanffy-Richards I: $Y_2 = \alpha\left(\frac{Y_1}{\alpha}\right)^{\frac{\ln(1 - e^{(-\beta T_2)})}{\ln(1 - e^{(-\beta T_1)})}}$

Polymorphic Von Bertalanffy-Richards II: $Y_2 = \alpha\left\{1 - \left[1 - \left(\frac{Y_1}{\alpha}\right)^{1-\beta}\right]^{\frac{T_2}{T_1}}\right\}^{\frac{1}{1-\beta}}$

Polymorphic Von Bertalanffy-Richards III: $Y_2 = \alpha\left\{1 + \left[\left(\frac{\alpha}{Y_1}\right)^{\beta} - 1\right] e^{[-\beta (T_2 - T_1)]}\right\}^{\frac{1}{\beta}}$

Anamorphic Von Bertalanffy-Richards I: $Y_2 = Y_1\left[\frac{1 - e^{-\beta T_2}}{1 - e^{-\beta T_1}}\right]^{\gamma}$

Table 4.3: Initial model fitting results for basal area and mean top height.

Equation:	Mean Square Error	
	G	MTH
Polymorphic Schumacher I	35.64	2.37
Polymorphic Schumacher II	34.22	1.5
Polymorphic Schumacher III	34.76	2.31
Anamorphic Schumacher I	42.06	4.07
Anamorphic Schumacher II	41.3	1.75
Polymorphic Hossfeld	38.24	1.51
Anamorphic Hossfeld	202.4	5.4
Anamorphic Weibull	40.61	1.75
Polymorphic Weibull I	32.79	Failed to converge
Polymorphic Weibull II	34.04	Failed to converge
Anamorphic Gompertz	39.89	1.79
Polymorphic Gompertz I	51.44	1.69
Polymorphic Gompertz II	33.36	1.47
Polymorphic Von Bertalanffy-Richards I	43.48	1.47
Polymorphic Von Bertalanffy-Richards II	Failed to converge	1.51
Polymorphic Von Bertalanffy-Richards III	Failed to converge	Failed to converge
Anamorphic Von Bertalanffy-Richards I	40.55	1.75

4.3.2.1 Calibration of the effect of elevation

Initial model fitting results showed a bias of increasing residuals for both MTH and basal area with increasing elevation, as was also found by Zhao (1999). To correct for this effect, elevation was incorporated into the four best initial models for both MTH and basal area.

Basal Area

The four models with the best initial fit for basal area were; Polymorphic Gompertz II, Polymorphic Schumacher II, Polymorphic Weibull I, and Polymorphic Weibull II as shown in Table 4.3. The effect of elevation was added into these models by altering the asymptotic parameter for each and results of model fit in terms of MSE are shown in Table 4.4. The model with the lowest MSE after the addition of the effect of elevation was altered further to see if different correction forms for the effect of elevation would

improve model fits. These included a linear correction for elevation for PSP's above 250 m, as was used by Zhao (1999), inclusion of the square of elevation, a probit correction for the effect of elevation, and a square of elevation combined with a linear correction for PSP's above 450 m elevation.

Table 4.4: Basal area model fitting results, after the effect of elevation is included in the asymptotic parameter.

Equation with elevation in asymptote	Basal Area MSE
Polymorphic Weibull I	failed to converge
Polymorphic Gompertz II	18.03
Polymorphic Weibull II	21.36
Polymorphic Schumacher II	18.38

The best model fit was found by adjusting the asymptotic parameter of the Polymorphic Gompertz II equation using elevation, elevation squared, and a dummy variable for elevations above 450m. The final model for basal area has a MSE of 17.02 and is written as:

$$G_2 = e^{\left(\ln(G_1)\cdot e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}+\frac{\left(\alpha_0+\alpha_1\cdot elev+\alpha_2\cdot elev^2+\alpha_3\cdot((elev-450)\cdot X)\right)}{10000}\cdot\left(1-e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right)\right)} \quad (4.31)$$

Where G_2 is the future basal area (m^2 / ha), G_1 is the initial basal area (m^2 / ha), T_1 is the initial stand age (years), T_2 is the final stand age (years), *elev* is the stand elevation (m), X is a binary indicator variable, $X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450, and β , γ , α_0 , α_1 , α_2 , and α_3 are parameters whose values are listed in table 4.5.

The parameter estimates of the final model of basal area are listed in table 4.5. The model showed no signs of bias for basal area / ha when plotted against prediction, age, elevation, and time increment (Figure 4.2).

Table 4.5: Parameters for basal area model (Equation 4.31), standard errors, and approximate 95% confidence limits calculated with 3666 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
β	0.1628	0.00201	0.1589	0.1668
γ	0.00261	0.00004	0.00253	0.00269
α_0	44797	255.1	44296.8	45297.2
α_1	-8.0659	1.4223	-10.8544	-5.2774
α_2	0.0491	0.00272	0.0437	0.0544
α_3	-37.8996	2.51	-42.8209	-32.9783

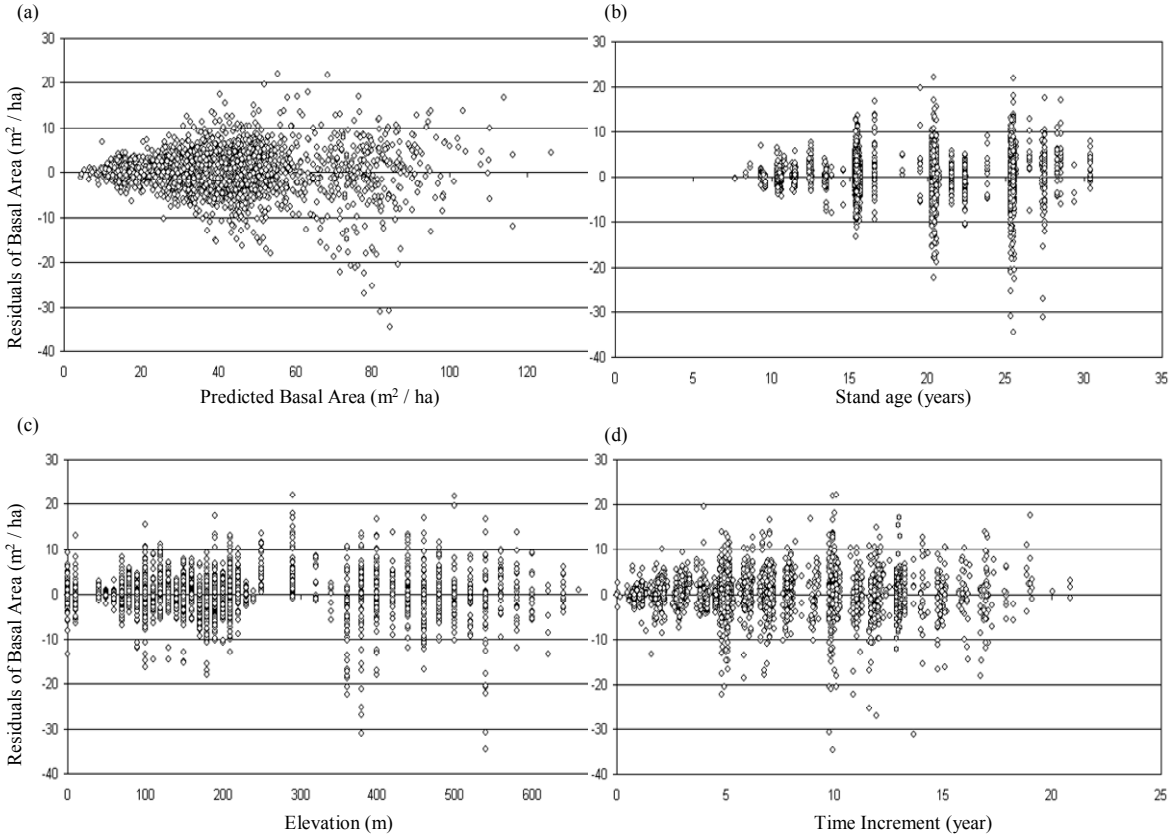


Figure 4.2: Fitting residual patterns of the final model for basal area: (a) residuals versus predicted, (b) residuals versus stand age, (c) residuals versus elevation, and (d) residuals versus time increment.

Mean Top Height

The four models with the best initial fit for mean top height where; Polymorphic Gompertz II, Polymorphic Von Bertalanffy-Richards II, Polymorphic Schumacher II, and the Polymorphic Hossfeld as shown in Table 4.3. The effect of elevation was added into these models by altering the asymptotic parameter for each and results of model fit in terms of MSE are shown in Table 4.6. The model with the lowest MSE after the addition of the effect of elevation was altered further. This was done to see if different correction forms for the effect of elevation would improve model fits.

Table 4.6: Mean top height model fitting results, including the effect of elevation in the asymptotic parameter.

Equation with elevation in asymptote	Mean Top Height MSE
Polymorphic Gompertz II	1.11
Polymorphic Von Bertalanffy-Richards II	1.10
Polymorphic Schumachar II	1.07
Polymorphic Hossfeld	1.09

The best model fit was found by adjusting the asymptotic parameter of the Polymorphic Schumacher II equation using elevation, elevation squared, and a dummy variable for elevations above 450m. The final model for mean top height had a MSE of 1.03 and is written as:

$$MTH_2 = e^{\ln(MTH_1)\cdot\left(\frac{T_1+\gamma}{T_2+\gamma}\right)^{\beta} + \frac{(\alpha_0+\alpha_1\cdot elev+\alpha_2\cdot elev^2+\alpha_3\cdot(elev-450)\cdot X)}{10000}\cdot\left(1-\left(\frac{T_1+\gamma}{T_2+\gamma}\right)^{\beta}\right)}$$

(4.32)

Where MTH_2 is the future mean top height (m), MTH_1 is the initial mean top height (m), T_1 is the initial stand age (years), T_2 is the final stand age (years), *elev* is the stand elevation (m), X is a binary indicator variable, $X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450, and β , γ , α_0, α_1, α_2, and α_3 are parameters whose values are listed in table 4.7.

The parameter estimates of the final model of mean top height are listed in table 4.7. The model showed no signs of bias for mean top height when plotted against prediction, age, elevation, and time increment (Figure 4.3).

Table 4.7: Parameters for mean top height model (Equation 4.32), standard errors, and approximate 95% confidence limits calculated with 3666 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
β	0.7613	0.0466	0.6699	0.8526
γ	4.4616	0.5125	3.4568	5.4664
α_0	44521.7	609.6	43326.4	45716.9
α_1	-7.3498	1.1851	-9.6733	-5.0264
α_2	0.0342	0.00255	0.0292	0.0392
α_3	-34.604	2.8022	-40.0982	-29.1098

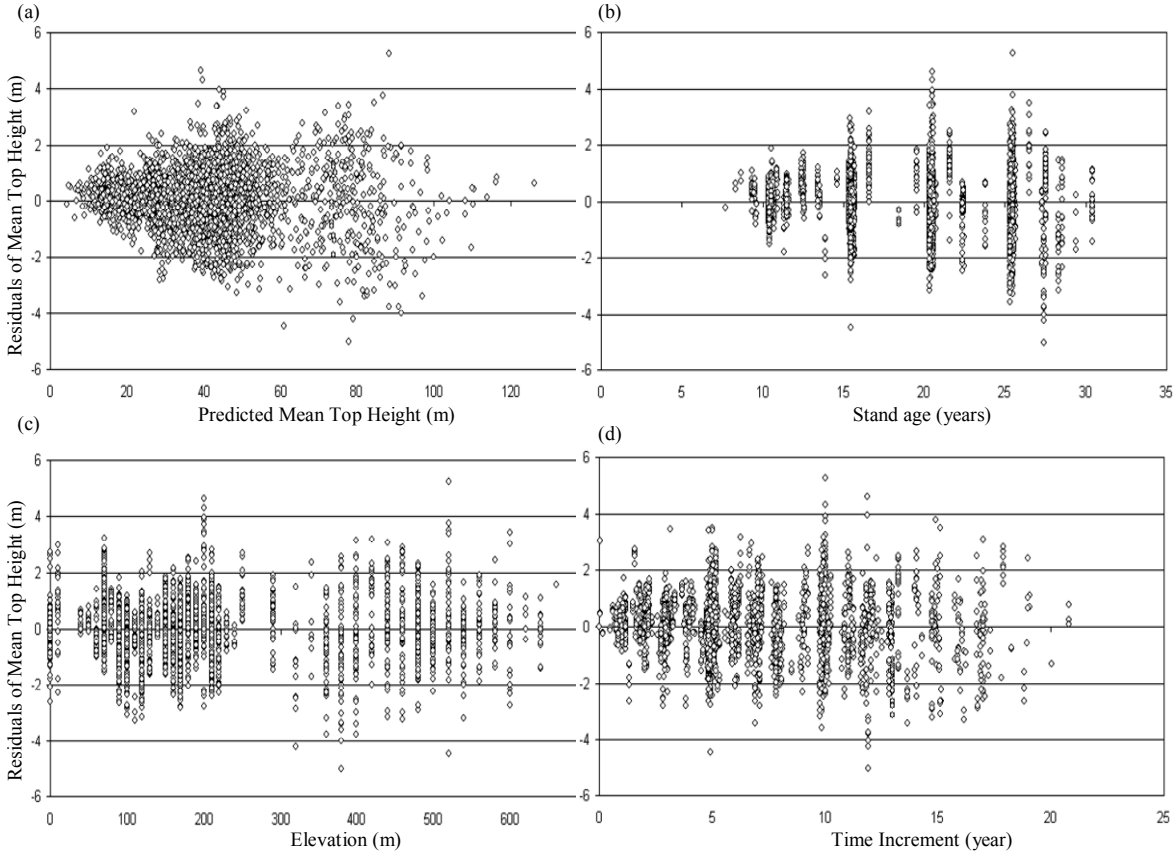


Figure 4.3: Fitting residuals patterns of final model for mean top height: (a) residuals versus predicted, (b) residuals versus stand age, (c) residuals versus elevation, and (d) residuals versus time increment.

4.3.3: Modelling maximum diameter and standard deviation of diameter

Models for maximum diameter and standard deviation of diameter were fitted by selecting the 4 model forms that had the best fit for basal area. These forms were also fitted to the data, incorporating the effects of altitude within the asymptotic parameter where appropriate. Results for the initial fitting of these models for maximum diameter and standard deviation of diameter are shown in Table 4.7.2.

Table 4.7.2: Mean square error for initial models of maximum diameter and standard deviation of diameter including the effect of elevation.

Equation with elevation in asymptote	Maximum Diameter	Standard Deviation of Diameter
Polymorphic Weibull I	5.13	.40
Polymorphic Gompertz II	5.38	.39
Polymorphic Weibull II	Failed to converge	.46
Polymorphic Schumacher II	5.28	.40

The best model fit for maximum diameter was found by adjusting the asymptotic parameter of the Polymorphic Weibull I equation using elevation squared, a dummy variable for elevations above 450m, and the inverse of initial stocking (Eq. 4.41). The best model fit for the standard deviation of diameter was similarly found by adjusting the asymptotic parameter for the Polymorphic Gompertz II equation using elevation, elevation squared and a dummy variable for elevations above 450m (Eq. 4.42). The final models for maximum diameter and standard deviation of diameter had MSE values of 4.62 and 0.39 respectively and are written as:

$$DMAX_2 = DMAX_1 \cdot e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)} + \left(\alpha_0 + \alpha_1 \cdot elev^2 + \alpha_2 \cdot \left(elev - 450\right)\cdot X + \alpha_3 \cdot \frac{1}{N_1}\right)\cdot\left(1 - e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)}\right)$$

(4.41)

$$DSTD_2 = e^{\left[\ln\left(DSTD_1\right)\cdot e^{\left(-\beta\cdot\left(T_2-T_1\right)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right]} \cdot e^{\left[\frac{\left(\alpha_0+\alpha_1 * elev^2+\alpha_2\cdot\left(elev-450\right)\cdot X\right)}{10000}\cdot\left(1-e^{\left(-\beta\cdot\left(T_2-T_1\right)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right)\right]}$$

(4.42)

Where $DMAX_2$ is the future maximum diameter (m), $DMAX_1$ is the initial maximum diameter (m), $DSTD_2$ is the future standard deviation of diameter (m), $DMAX_1$ is the initial standard deviation of diameter (m),T_1 is the initial stand age (years), T_2 is the final stand age (years), *elev* is the stand elevation (m), X is a binary indicator variable, $X = 0$ if elevation < 450, N1 is the initial stocking (stems / ha) and $X = 1$ if elevation ≥ 450, and β , γ , α_0 , α_1 , α_2 , and α_3 are parameters whose values are listed in table 4.8 (for maximum diameter eq. 4.41) and table 4.9 (for standard deviation of diameter eq. 4.42).

Tabeles 4.8, and 4.9 show parameter estimates. Figures 4.4, and 4.5 show residual patterns with little apparent bias against main modelling variables. For projection lengths of up to 20 years 95% of residuals for maximum diameter projections were within ± 4.3cm, while 95% of the residuals for standard deviation of diameter were within ± 1.3 cm.

Table 4.8: Parameter values for maximum diameter model (Equation 4.41). Also shown are standard errors, and approximate 95% confidence limits for each parameter. Statistical values are calculated with 3666 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
γ	0.3377	0.0378	0.2637	0.4118
β	0.4128	0.0333	0.3476	0.478
α_0	66.0643	3.1822	59.8252	72.3034
α_1	0.000165	0.000011	0.000144	0.000186
α_2	-0.1967	0.018	-0.232	-0.1614
α_3	8257.1	699	6886.6	9627.6

Table 4.9: Parameter values for standard deviation of diameter model (Equation 4.42). Also shown are standard errors, and approximate 95% confidence limits for each parameter. Statistical values are calculated with 3666 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
β	0.0725	0.00178	0.069	0.076
γ	0.000458	0.000061	0.000337	0.000579
α_0	24257.6	296.6	23676.1	24839.1
α_1	0.0194	0.000979	0.0175	0.0213
α_2	-33.2487	3.1668	-39.4576	-27.0398

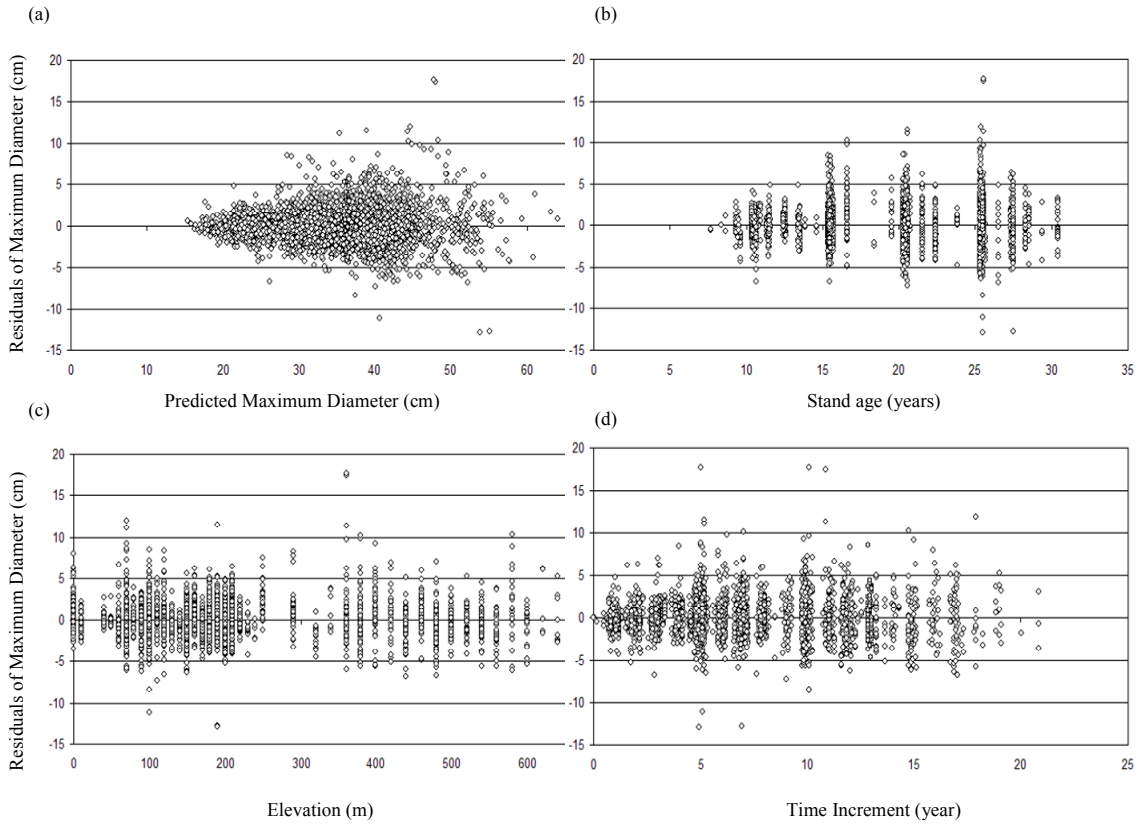


Figure 4.4: Residual patterns of final model for maximum diameter: (a) residuals versus predicted, (b) residuals versus stand age, (c) residuals versus elevation, and (d) residuals versus time increment.

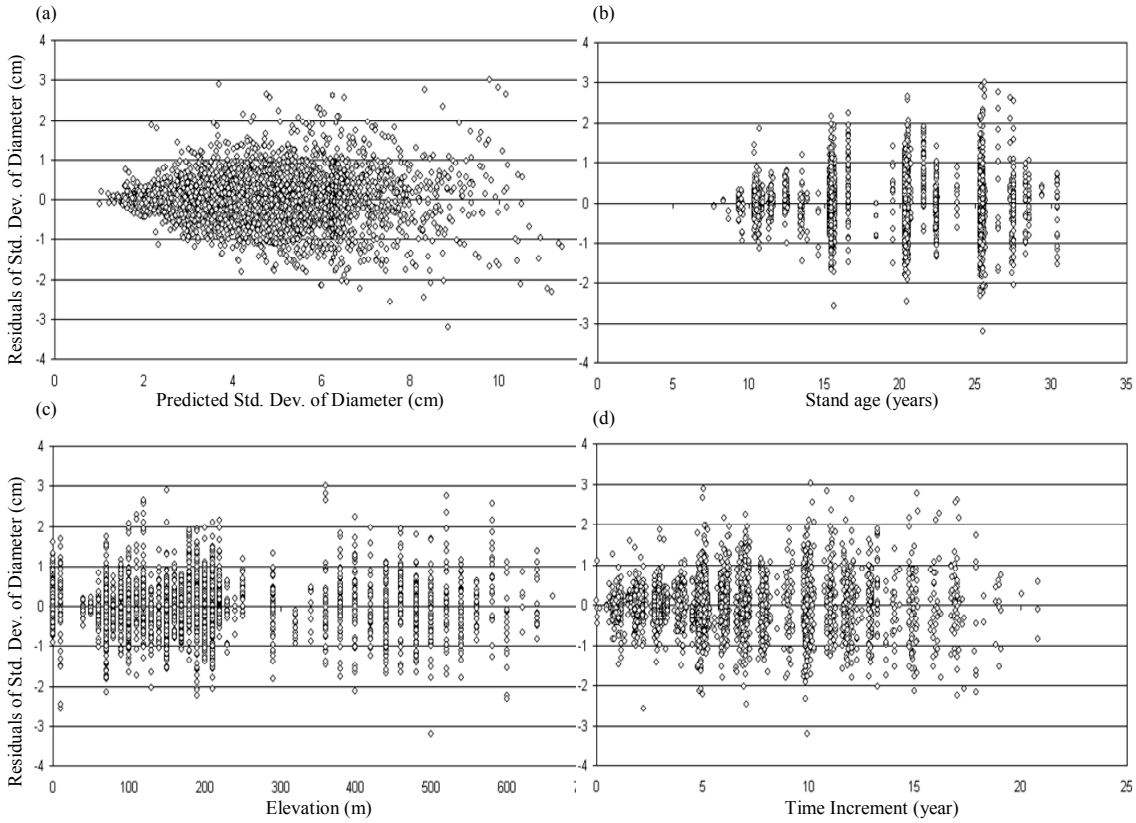


Figure 4.5: Residuals patterns of final model for standard deviation of diameter: (a) residuals versus predicted, (b) residuals versus stand age, (c) residuals versus elevation, and (d) residuals versus time increment.

4.3.4: Modelling Stems per Hectare

Mortality Modelling Methods

A set of six equations which are variants of those listed by Clutter *et al.* (1983) and those used directly by Zhao (1999) were used as the basic equation forms for projecting stocking (Table 4.51). These equations were initially fitted to the modelling data and the results are listed in table 4.10. As equation 1 had the smallest MSE it was chosen as the basic form of difference equation for further analysis. Methods used to improve the fit of this equation were; (1) using a two step logistic modelling approach, (2) the addition of maximum wind speed over the observation period to the basic equation, and (3) using a mortality severity index to filter the modelling data, which is intended to create a model that is unbiased with respect to large uncommon mortality events.

Table 4.10: Difference equation forms for projection of stocking.

Difference Equation Forms	Parameter Estimates		Std.Error	Model MSE
1: $N_2 = \left(N_1^{\,c} + \frac{a}{100000} \cdot \left(T_2^{\,b} - T_1^{\,b} \right) \right)^{\left(1/c\right)}$	*a*	0.00316	0.00146	3831
	b	2.5116	0.1369	
	c	-1.1046	0.0461	
2: $N_2 = \left(\frac{1}{\sqrt{N_1}} + \frac{a}{100000} \cdot \left(T_2^{\,b} - T_1^{\,b} \right) \right)^{-2}$	*a*	0.0137	0.00587	4066
	b	3.1076	0.1284	
3. $N_2 = \left(\frac{1}{\sqrt{N_1}} + a \cdot \left(\left(\frac{T_2}{100} \right)^2 - \left(\frac{T_1}{100} \right)^2 \right) \right)^{-2}$	*a*	0.0571	0.000975	4181
4. $N_2 = N_1 \cdot e^{b \cdot (T_2 - T_1)}$	*b*	-0.0110	0.000233	5429
5. $N_2 = N_1 \cdot \left(\frac{T_2}{T_1} \right)^a \cdot e^{b \cdot (T_2 - T_1)}$	*a*	0.3930	0.0163	4690
	b	-0.0355	0.00105	
6. $N_2 = N_1 \cdot e^{a \cdot \left(T_2^{\,b} - T_1^{\,b} \right)}$	*a*	-2.17E-6	9.992E-7	4701
	b	3.5101	0.1387	

Mortality Modelling Method 1

The first mortality model attempted to improve upon the original fit of equation 1 (Table 4.10) followed the methods outlined in Woollons (1998) for a two step regression model that uses both logistic and non-linear approaches. This model was chosen because the event of stem death occurring or not in a particular permanent sample plot over a given period of time is a binomial outcome (mortality or no mortality), which can be modelled using a logistic regression (Hosmer and Lemeshow, 1989). A two step method suggested by Woollons (1998) to model mortality in even aged forests first uses a logistic

regression to predict the probability of stems/ha death. A mortality equation is then built, but only utilising data where death has occurred over the measurement interval (in this case Eq. 1 from Table 4.10 was used as the basic model form). Estimates from the second model are then reduced by a factor, equivalent to the probability of death occurring, acquired from the logistic equation.

The equations used for modelling the probability of mortality and adjusting the final number of stems per hectare based on the probability of death have been outlined in Woollons (1998) as follows:

If you denote the event of Mortality or (0,1) as Y, let X_1, X_2…, X_n be a set of explanatory variables (in this application, N, T and perhaps S, site index) and let 'p' be the probability of stem death occurring over a period.
Then we have:

$$\text{logit}(p) = \log\left(\frac{p}{(1-p)}\right) = \beta_0 + \beta_1 X_1 + \ldots \beta_n X_n$$

or

$$p = 1/\left(1 + \exp\left(-\left(\beta_0 + \beta_1 X_1 + \ldots \beta_n X_n\right)\right)\right) \qquad (4.51)$$

(2) Model mortality, only using the data which exhibits some mortality, through a projection model of basic form Eq (4.51).

(3) Estimate live stems / ha at time T_2 through:

$$N_{adj2} = N_1 - p \cdot \left(N_1 - N_2\right)$$

or

$$\frac{N_{adj2}}{N_1} = p \cdot \left(\frac{N_2}{N_1}\right) + \left(1 - p\right) \qquad (4.52)$$

where N_{adj2} = (adjusted)live stems/ha at time T_2; N_1=live stems/ha at the beginning of the period; N_2 = live stems / ha at time T_2 (using Eq. 1 from table 4.10 fitted to the modelling data set filtered to include only observations where mortality occurred), estimated by a model of form Eq. (1); p = probability that Y = 1 (death has occurred).

For the second step in the logistic modelling procedure, the dataset was reduced down to measurement intervals that occurred most commonly within the original modelling dataset so that the final model form would not need to include time interval as a predictor variable. The mode re-measurement interval in the fitting data set was five years (Table 4.10)

Table 4.11: Re- measurement interval descriptive statistics.

Mean	6.34
Standard Error	0.063
Median	5.09
Mode	5
Standard Deviation	4.01
Sample Variance	16.13
Kurtosis	0.24
Skewness	0.82
Range	20.82
Minimum	0
Maximum	20.82
Sum	25092.58
Frequency	3957

Mortality Modelling Method 2

The next modelling approach was to add the environmental variable of maximum wind speed alone and the interaction of wind speed and mean top height to Equation 1 (Table 4.10) to see if this would improve model fit. Maximum wind speed over the measurement interval was obtained by measurements from Christchurch Airport (S43.5°, E172.55°, elevation 37m).

Mortality Modelling Method 3

The final mortality model was made in consultation with managers at SPBL to ensure that it did not include any mortality events that were considered catastrophic wind events that may not occur regularly over a rotation. Catastrophic mortality events artificially increase estimates of mortality in all plots and cause an underestimation of standing volume.

The guide used for selecting a dataset to model mortality was based on the -3/2 power law (Yoda *et al.* 1963).

$$W = C \cdot N^{\left(\frac{-3}{2}\right)} \tag{4.53}$$

Or

$$\log W = -\left(\frac{3}{2}\right) \cdot \log N + \log C \tag{4.54}$$

Where:

W = plant weight in kg, N = plant density (stems/ha), C= constant

This law states that the slope of the relationship between log(stocking) and the log(stem mass) is linearly related and has a slope of -3/2. This law is based on the geometrical overcrowding of growing space and describes mortality caused by competition for space.

An index of the severity of a mortality event was based on the -3/2 relationship that took into account the distance a particular stand was from this -3/2 line (where stands that were farthest from the line would indicate a high mortality event) and the index also took into account the slope of a particular stand over a measurement interval with respect to this log stocking vs. log stem mass relationship in the -3/2 power law. A stand that displayed flatter slopes between the first and second measurement period (T_1 and T_2) of the log stocking vs. the log stem mass relationship, was assumed to be a stand that had a severe mortality event. A flatter slope between the first and second measurement would indicate a stand that had significant drop in stocking without an accompanying increase in stem mass, which occur in pulse mortality events.

The mortality severity index is written as:

Mort Index = (proportional distance the measurement of stocking at T_2 is from -3/2 line)*1/(slope of log log relationship between measurements of stocking at an initial and final age (T_1 and T_2).

The proportional distance is defined by the distance T_2 (an observation of stocking at some final age) is from the -3/2 line in relation to the distance between -3/2 and a parallel line passing through the origin (Figure 4.6).

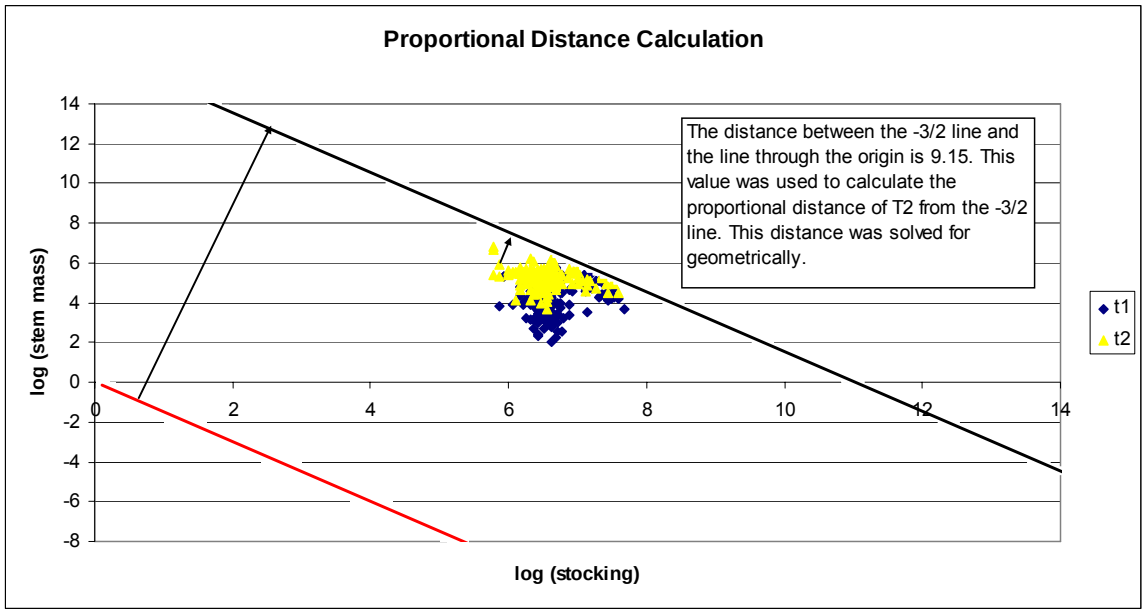


Figure 4.6: Proportional distance calculation used in mortality severity index.

The distance from the point T_2 to the -3/2 line was calculated using the point-line distance formula (Eq. 4.55):

$$d = \frac{\left|(x_2 - x_1)(y_1 - y_0) - (x_1 - x_0)(y_2 - y_1)\right|}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} \qquad (4.55)$$

Where: (x_0, y_0) is a point in space some perpendicular distance d from the line containing points (x_1, y_1) and (x_2, y_2) (Figure 4.7)

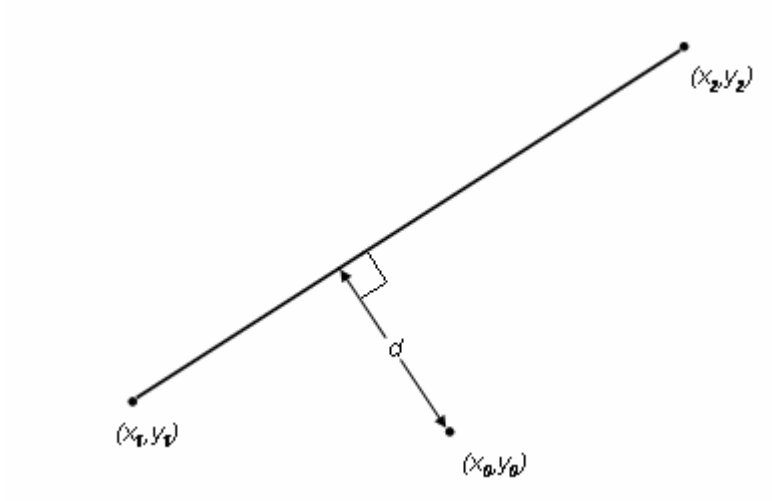


Figure 4.7: Point line distance formula (Eq. 4.55) reference figure.

Percentages of the dataset were selected based on a cut-off of this severity index, and was intended to describe the average condition for mortality on the SPBL estate. Cut-off levels were selected to pick out the most severe mortality events based on this index. Cut-off levels were selected to leave 80%, 60%, and 40% of the original dataset (Figure 4.8). Four models were created by fitting these three datasets to equation 1 from Table 4.10.

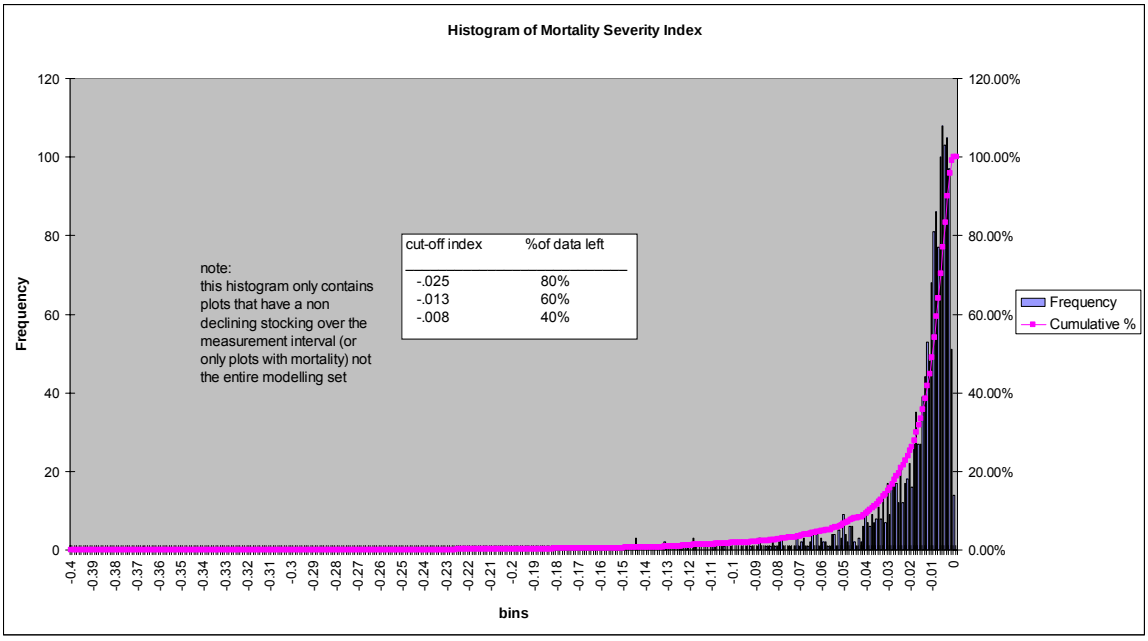


Figure 4.8: Histogram of mortality severity index cut off values.

Preliminary models created from the three datasets were given to the managers at SPBL for inspection. The final model chosen was the one that described mortality levels that

fitted their view from management experience of what "non-catastrophic" mortality was. This was necessary because they did not wish models to include devastating wind throw events in calculations of expected mortality from year to year. The Canterbury region suffers wind throw catastrophes every one or two decades, and these require specific study that was beyond the scope of this thesis.

Mortality Modelling Results

Resulting MSE from the two-step logistic regression process (method 1) yielded a MSE of 4374. As this MSE was higher than the original fit of Equation 1 in Table 4.10, the logistic two step approach to modelling mortality was not used as the final model.

The results of adding the interaction of wind speed and mean top height to Equation 1 (Table 4.10) (method 2) showed a modest improvement in MSE of 3.6 % (MSE = 3693). This improvement in predictive power was not seen as being significant enough in scale to warrant the inclusion of maximum wind speed into the mortality equation. Wind data is awkward to use in a modelling context, as it is so variable throughout a given time period, and is not measured accurately at PSPs. Data extrapolation methods give good measurements of average values over long periods, but may not describe pulse wind events that would dictate mortality. Including wind as a variable in mortality models would only work well for past projections where accurate measurements of wind speed have been measured and may be of limited value for future projections as wind speeds input into models would have to be average values. Future wind speed in a given area would be hard to predict accurately enough to make its inclusion worth while.

In the final approach (method 3) of segregating the modelling dataset with respect the mortality severity index, managers at SPBL selected a data set cut off value of 80% as being most representative of mortality on their estate. The final model selected (Eq. 4.56) was fitted using Equation 1 (Table 4.10) and produced a MSE of 4128. Parameters for the final mortality model (Eq. 4.56) are listed in Table 4.12. As this model included a reduced number of observations, MSE cannot be compared with those outlined previously. Projections with this model were not as precise as those of MTH and basal area, as shown with residual patterns in Figure 4.9. Little bias was apparent, however.

Ninety percent (>5% and <95%) of residuals were within ± 75 stems / ha for projection lengths of up to 21 years.

$$N_2 = \left(N_1^{-1.0169} + \frac{0.00740}{100000} \cdot \left(T_2^{2.4358} - T_1^{2.4358} \right) \right)^{\left(1/-1.0169\right)} \qquad (4.56)$$

Where N_2 is the future stocking (stems / ha), N_1 is the initial stocking (stems / ha), T_1 is the initial stand age (years), T_2 is the final stand age (years).

Table 4.12: Stocking model parameters for equation 4.56, standard errors, and approximate 95% confidence limits calculated with 3280 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
a	0.0074	0.00357	0.000409	0.0144
b	2.4358	0.1414	2.1585	2.7131
c	-1.0169	0.0478	-1.1106	-0.9231

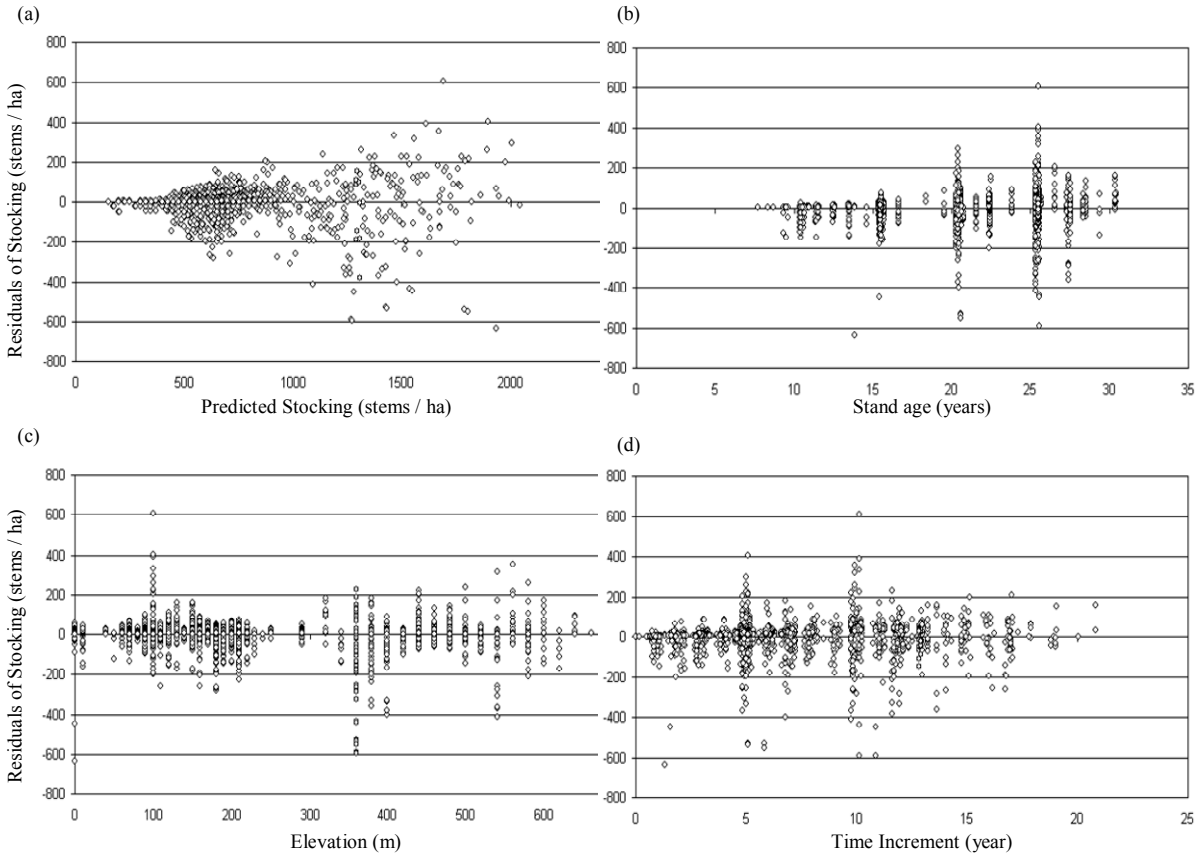


Figure 4.9: Residuals patterns of final model for stocking: (a) residuals versus predicted, (b) residuals versus stand age, (c) residuals versus elevation, and (d) residuals versus time increment.

4.4: Testing Model Parameters

Final model parameters were tested against an autocorrelation-free dataset to see if they were significantly different from zero. An autocorrelation-free dataset was prepared from the original modelling data by randomly selecting one observation from each of the 749 plots. Parameters for the models of mean top height, basal area, maximum diameter, and standard deviation of diameter were all found to be significant at 95% confidence level (Tables 4.13, 4.14, 4.15, and 4.16).

Table 4.13: Basal area model parameters for equation 4.31, standard errors, and approximate 95% confidence limits tested with an autocorrelation-free dataset with 748 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
β	0.1613	0.0044	0.1526	0.1699
γ	0.00266	0.000087	0.00249	0.00283
α_0	45450.3	607.8	44257	46643.5
α_1	-11.2773	3.5706	-18.2869	-4.2676
α_2	0.0573	0.00686	0.0438	0.0707
α_3	-49.4855	6.0197	-61.3033	-37.6677

Table 4.14: Mean top height model parameters for equation 4.32, standard errors, and approximate 95% confidence limits tested with an autocorrelation-free dataset with 748 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
β	6.3018	1.3502	3.6512	8.9524
γ	0.9334	0.1218	0.6942	1.1726
α_0	42511.3	1034.2	40481	44541.5
α_1	-4.6719	2.3254	-9.2372	-0.1067
α_2	0.0293	0.005	0.0195	0.0391
α_3	-34.0153	5.3091	-44.438	-23.5926

Table 4.15: Maximum diameter model parameters for equation 4.41, standard errors, and approximate 95% confidence limits tested with an autocorrelation-free dataset with 748 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
γ	0.3796	0.0878	0.2072	0.552
β	0.354	0.0675	0.2215	0.4866
α_0	76.0848	10.0834	56.2891	95.8805
α_1	0.000215	0.000036	0.000144	0.000287
α_2	-0.2675	0.052	-0.3696	-0.1653
α_3	8029.7	1686	4719.8	11339.5

Table 4.16: Standard deviation of diameter model parameters for equation 4.42, standard errors, and approximate 95% confidence limits tested with an autocorrelation-free dataset with 748 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
β	0.0639	0.00348	0.0571	0.0708
γ	0.000446	0.000114	0.000223	0.00067
α_0	26070.7	935.7	24233.8	27907.6
α_1	0.0137	0.00251	0.00873	0.0186
α_2	-18.4663	7.5379	-33.2646	-3.668

Two out of three parameters for the mortality model were found to be significant at the 95% confidence level (Tables 4.17). Although the asymptotic parameter for Equation 1

was found to be non-significant when tested against an autocorrelation free dataset, it was chosen as the final mortality equation. A simpler mortality model (Equation 3, Table 4.10) was found to have significant parameters when tested against and autocorrelation free data set. This model however, had a worse initial fit (Table 4.10), and was found to also have a higher MSE than Equation 1, when tested against the validation data set.

Table 4.17: Mortality model parameters for equation 4.56, standard errors, and approximate 95% confidence limits tested with an autocorrelation-free dataset.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
a	0.000217	0.000255	-0.00028	0.000718
b	3.9103	0.3249	3.2724	4.5482
c	-0.785	0.0961	-0.9737	-0.5963

4.6: Model Validation and Comparison with CanSPBL(1.0)

Validation of CanSPBL(1.2) and CanSPBL1 with the same validation data set presented previously in this chapter, produce model fitting statistics and residual distribution statistics listed in Tables (4.18) and Table (4.19) respectively. CanSPBL(1.2) showed improvements for all model components of 4 to 46% after being updated with a more current and complete dataset covering a better range of elevation and stand ages across the SPBL estate. CanSPBL(1.2) also showed less bias in prediction of standard deviation of diameter, and mean top height (Table 4.18). Residual distributions for the updated model showed a better fit overall for skewness and kurtosis, while the hypothesis of normality of residuals was not rejected for both models with the Kolmogorov-Smirnov test (Table 4.19).

Table 4.18: Comparison of validation model statistics between CanSPBL, and CanSPBL(1.2).Statistics shown include Mean Square Error (MSE), Average Model Bias (AMB), Model Efficiency Factor (EF), and percent difference of MSE (% Difference of MSE) n = 969.

Model		Model Fitting Statistics			
		MSE	AMB	EF	MSE % Difference
CanSPBL(1.0)					
	Stocking	2247.22	-6.89	0.95	4.39
	Basal Area	18.64	-0.54	0.96	4.04
	Mean Top Height	1.77	0.48	0.94	46.39
	Maximum Diameter	6.70	-0.23	0.92	17.85
	Standard Deviation of Diameter	0.71	0.21	0.82	28.60
CanSPBL2(1.2)					
	Stocking	2150.58	6.52	0.96	-4.39
	Basal Area	17.91	0.55	0.96	-4.04
	Mean Top Height	1.11	0.08	0.96	-46.39
	Maximum Diameter	5.59	0.27	0.93	-17.85
	Standard Deviation of Diameter	0.53	0.04	0.86	-28.60

Table 4.19: Model normality distribution statistics for CanSPBL, and CanSPBL(1.2) in terms of skewness, kurtosis, and p-value from the Kolmogorov-Smirnov test for normality n = 969.

Model		Residual Distribution Statistics		
		Skewness	Kurtosis	P-value
CanSPBL(1.0)				
	Stocking	-2.86	15.02	< 0.01
	Basal Area	-0.01	3.96	< 0.01
	Mean Top Height	0.74	1.33	< 0.01
	Maximum Diameter	1.20	6.25	< 0.01
	Standard Deviation of Diameter	1.79	10.55	< 0.01
CanSPBL2(1.2)				
	Stocking	-1.01	11.08	< 0.01
	Basal Area	0.42	3.75	< 0.01
	Mean Top Height	0.12	0.95	< 0.01
	Maximum Diameter	1.15	4.72	< 0.01
	Standard Deviation of Diameter	1.22	7.33	< 0.01

4.7: Discussion

Refitting the equations that predict basal area, mean top height, standard deviation of diameter, and stocking in the growth and yield model CanSPBL(1.0) with a more current data set that better covered a range of elevations and stand ages on the Selwyn estate showed a large significant improvement in model accuracy (Table 4.18).

The effects of elevation incorporated into these models were somewhat complicated and indicate elevation may be a proxy for physiological factors that effect growth or variations in site factors. Model accuracy may be increased further by incorporating physiological factors such as root zone water balance over the measurement interval, or possibly the amount of used photosynthetically active radiation. Future research in this area should focus on trying to improve model accuracy and robustness by incorporating process-based approaches that will simplify the form of elevation in these models and make them applicable to changing climatic and site conditions.

4.8: Conclusions

Model components of mean top height, basal area per hectare, stems per hectare, and diameter distribution were developed by using non linear least squares regression techniques to select appropriate equation forms. Explanatory variables to improve model prediction were tested and incorporated into models where appropriate. The effect of elevation was added to models of mean top height, basal area, and diameter distribution. A polymorphic Gompertz equation displayed the best fit for basal area, while a polymorphic Schumacher equation displayed the best fit for mean top height. The diameter distribution model of maximum diameter displayed the best fit with a polymorphic Weibull equation, while the standard deviation of diameter model fit best with a polymorphic Gompertz equation.

Many different modelling approaches were tried to model the number stems per hectare at the end of a simulation period. Logistic two step modelling approaches did not show an improvement over a standard 3 parameter difference equation, and the inclusion of

maximum wind speed over the measurement interval resulted in only a modest improvement of 3.6 %. A system of modelling mortality with a three parameter difference equation, using a dataset filtered by a mortality severity index was chosen as the final modelling approach. This method reduced the modelling dataset to only include mortality events that were characteristic of the manager's view of average mortality. A mortality severity index based on the -3/2 power law was used as a basis to filter the modelling data set for mortality and a final model was built with 80% of the original mortality records under the recommendation of managers at SPBL. Residual mean square error for the model of stocking was decreased by 4.39 % by refitting equations with a more current dataset (equation 4.57). Ninety percent (>2.5% and <97.5%) of residuals were within ± 129 stems / ha for projection lengths of up to 21 years. Future stocking is dependent on the current stocking (stems / ha), initial stand age (years), and final stand age (years).

$$N_2 = \left(N_1^{-1.0169} + \frac{0.00740}{100000} \cdot \left(T_2^{2.4358} - T_1^{2.4358} \right) \right)^{\left(1/-1.0169\right)} \qquad (4.57)$$

All model parameters were tested against an auto-correlation free dataset for significance. Tests showed that all but one parameter was significant at a 95% confidence level. The model with the insignificant parameter was kept as a final equation form after validating the model to see if in fact that form made the most accurate predictions against a validation dataset.

The entire model was validated against an independent data set and compared to a the original CanSPBL(1). CanSPBL(1.2) showed improvements in MSE of 4 to 46% after being updated with a more current and complete dataset that covered a better range of elevation and stand ages across the SPBL estate. Residual distributions for the updated model showed a better fit overall for skewness and kurtosis, with the exception of the model for mean top height where skewness is slightly higher for the updated model. The hypothesis of normality of residuals was not rejected for both models with the Kolmogorov-Smirnov test.

Residual mean square error for the model of basal area (equation 4.58) was reduced by 4.04% from the original CanSPBL(1.0). Ninety five percent (>2.5% and <97.5%) of residuals were within ± 8.31 m^2 / ha for projection lengths of up to 21 years. Future basal area is dependent on the current stand basal area (m^2 / ha), initial stand age (years), final stand age (years), elevation (m), elevation squared (m^2), and a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450).

$$G_2 = e^{\left(\ln(G_1)\cdot e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}+\frac{\left(\alpha_0+\alpha_1\cdot elev+\alpha_2\cdot elev^2+\alpha_3\cdot((elev-450)\cdot X)\right)}{10000}\cdot\left(1-e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right)\right)}$$

(4.58)

Residual mean square error for the model of mean top height (equation 4.59) was reduced by 46.39 % from the original CanSPBL(1.0). Ninety five percent (>2.5% and <97.5%) of residuals were within ± 2.07 m for projection lengths of up to 21 years. Future mean top height is dependent on the current mean top height (m), initial stand age (years), final stand age (years), elevation (m), elevation squared (m^2), and a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450).

$$MTH_2 = e^{\ln(MTH_1)\cdot\left(\frac{T_1+\gamma}{T_2+\gamma}\right)^{\beta}+\frac{\left(\alpha_0+\alpha_1\cdot elev+\alpha_2\cdot elev^2+\alpha_3\cdot(elev-450)\cdot X\right)}{10000}\cdot\left(1-\left(\frac{T_1+\gamma}{T_2+\gamma}\right)^{\beta}\right)}$$

(4.59)

Residual mean square error for the model of maximum diameter (equation 4.60) was reduced by 17.85 % from the original CanSPBL(1.0). For projection lengths of up to 21 years 95% of the residuals for maximum diameter projection were within ± 4.30 cm. Future maximum diameter is dependent on the current maximum diameter (cm), initial stand age (years), final stand age (years), elevation squared (m^2), and a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450).

$$DMAX_2 = DMAX_1\cdot e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)}+\left(\alpha_0+\alpha_1\cdot elev^2+\alpha_2\cdot(elev-450)\cdot X+\alpha_3\cdot\frac{1}{N_1}\right)\cdot\left(1-e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)}\right)$$

(4.60)

Residual mean square error for the model of standard deviation of diameter (equation 4.61) was reduced by 28.60 % from the original CanSPBL(1.0). For projection lengths of up to 21 years 95% of the residuals for standard deviation of diameter projections were within ± 1.24 cm. Future standard deviation of diameter is dependent on the current standard deviation of diameter (cm), initial stand age (years), final stand age (years), elevation squared (m^2), and a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450).

$$DSTD_2 = e^{\left[\ln(DSTD_1)\cdot e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right]}\cdot e^{\left[\frac{\left(\alpha_0+\alpha_1*elev^2+\alpha_2\cdot(elev-450)\cdot X\right)}{10000}\cdot\left(1-e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right)\right]}$$

(4.61)

CHAPTER 5

MODELLING LEAF AREA INDEX

5.1 Introduction

Leaf area index (LAI) is a critical variable in monitoring and modelling forest condition and growth. It is important for a number of physiological processes such as photosynthesis, transpiration, evaporation (Pierce and Running, 1988) and net primary production (Monteith, 1972). As a result leaf area is an important input in many ecosystem models that simulate carbon growth and hydrological cycles (Gower *et al.* 1999).

Patterns in leaf area development over time in vegetative canopies have a similar but distinctive pattern depending on species composition and resources available for growth. Foliage mass has been shown to increase with age until an equilibrium level is reached in both *Pinus radiata* (Madgwick *et al.* 1977) and *Pseudotsuga mensiesii* (Long and Smith 1984). A slightly different pattern in canopy development has been found in *Pinus sylvestris* stands which increase to a maximum over time and then decline sharply. This pattern of decline in foliage growth is thought to be due to competition, immobilisation of nutrients, and increased mutual shading (Kuuluvainen, 1991). In physiological studies of barley and other species canopy growth has also been shown to depend on site characteristics such as the availability of soil nitrogen (caused by the ratio of photosynthesis per gram of leaf) (Chapin *et al.,* 1988; Chapin, 1991; Evans, 1989; Field and Mooney,1986).

LAI can be measured directly through destructive sampling, or indirectly using optical techniques and stand site variables. Estimates of LAI have been correlated with individual tree and stand-level variables. In *Eucalyptus nitens*, individual tree leaf area has been shown to decrease with increasing initial stand stocking (Pinkard and Nelson,

2003). Pichler *et al.*(2001), have also shown a correlation of hemispherical photo measurements of LAI and diameter increment and terminal shoot length. Waring *et al.* (1982) suggest that there is a strong relationship between leaf area index and the sapwood area at the base of the live crown of a tree. Pinkard and Neilsen (2003) found that stand stocking was not significantly related to stand LAI. This was one of the few studies that related stand measurements to stand LAI and involved very small datasets because of the high cost of measuring LAI destructively. In contrast, the study reported here profited from relatively new digital image analysis techniques for measuring LAI, thus the dataset obtained for modelling was larger than those reported in previous studies.

The general objectives of this study were to derive two relatively simple models for predicting one sided canopy leaf area index (LAI) of radiata pine stands growing on the Canterbury Plains of New Zealand. The parameters of the models were hypothesised to be correlated with (1) elevation as a surrogate for temperature and mean annual rainfall; (2) slope and aspect, as southerly aspects are known to be wetter and therefore more productive in the Canterbury region; (3) basal area and mean top height. Mean top height is defined in New Zealand as the predicted height of a tree with the quadradic mean diameter of the 100 largest diameter trees in a one hectare area (Burkhart and Tennent, 1977). (4) An additional objective was to determine if parameters of the model were significantly related to stand density.

5.2 Methods

Data Collection

All data were collected from the Selwyn Plantation Board Ltd. (SPBL) estate located on the Canterbury Plains of south-eastern New Zealand (latitude range -43.77 to -43.33 and longitude range 171.72 to 172.71 east longitude). The SPBL estate covers 3 forest types of plains, hills, and coastal sands with elevations ranging from 2 – 600 metres above sea level. Distinctions between these forest types are due to both elevation and soil type. With increasing distance from the sea soils underlying forests in Canterbury change from coastal sands, shallow and dry floodplains soils, to deep wet LOESS hill soils (Zhao, 1999). Average temperatures are 10.9 degrees for plains and coastal forests, and 11.6 degrees for forests growing on the hills. Annual rainfall increases from 600 to 1100mm

for coastal-plains and hills forest respectively. The silvicultural strategy adopted by SPBL targets a final crop stocking from 400 to 650 stems per hectare, where stands are thinned around age 10. Stands are pruned in a two lift process to a final canopy height of 2.5 meters. This silvicultural strategy is applied to the entire estate and varies little between stands.

Hemispherical photos were taken in a sample of inventory plots that were measured in SPBL's 2004 summer inventory. Inventory measurements included tree heights, diameters at breast height (1.4m), and stand stocking. Plots were sampled to reflect a range of age classes, stand stockings, and site conditions. Photographs were taken from the forest floor looking skyward with a 180 degree hemispherical (fisheye) lens (Figure 5.1). The camera was located at the centre of each permanent sample plot and geographic north was marked in the photograph. Images were captured with a Nikon fisheye converter FC-E9 0.2x lens, mounted on a Nikon coolpix 5700 camera. An ocular estimate was recorded of percent ground cover in the forest that was not part of the overstorey canopy at each plot where a photograph was taken. The percent of understorey was recorded as it could have been confused with canopy elements during digital analysis, which may artificially increase estimates of LAI of the overstorey species. This estimate did not include branches and bark of overstorey trees. Estimates included other understorey species, weed overstorey species, and dead and downed wood, such as fallen trees and pruned branches. Along with an estimate of the percent of understorey ground cover, an ocular estimate of height relative to canopy height was also recorded.

The images were digitally analysed using the Gap Light Analyser software (Frazer *et al.,* 1999). Digital images were imported into Gap Light Analyser (GLA) in the form of bitmaps, which were then digitally analysed. Fisheye photography produces circular images that record the size, shape, and location of gaps in the forest overstorey. Image processing involves the transformation of image pixel positions into angular coordinates, the division of pixel intensities into sky and non-sky classes, and the computation of sky-brightness distributions. These data were combined to produce estimates of leaf area index (Chazdon and Field, 1987; Becker *et al.,* 1989; Rich, 1990; ter Steege, 1993; Canham 1995).

Figure 5.1: Sample of hemispherical photograph used for image analysis and calculation of LAI.

Inputs to GLA were magnetic north, slope (deg), elevation (m), aspect (deg), latitude and longitude. Plot coordinates provided by SPBL were converted from New Zealand map grid to a standard coordinate system. Slope, aspect, and elevation were derived in ArcView GIS (2000) with the Geo analysis tool pack. All topographic lines of the study area were spatially joined together to make a continuous cover (Land Information New Zealand, 2004). After joining was completed, slopes, aspects, and elevations were derived. To convert these values from a grid to a shape file, where they could be linked with plot coordinate information, the aspects and slopes had to be converted into categorical variables. Aspect was categorised into north, south, east, west, northwest, northeast, southwest, and southeast. Slopes were put into 5 degree classes (0-5, 5-10 etc.). The slope input into GLA was taken as the midpoint between these class variables, for example, the slope of a plot that was in the 0-5 degree class was input as 2.5 into GLA.

Modelling Data

Modelling data comprised LAI measurements, mensuration measurements, and available soil and climatic databases. Plots included in this study were a sample of 141 permanent

sample plots (PSP's) that were 0.04 ha in size, all located within the SPBL estate. Plots were chosen to represent a range of stand age, stocking, and elevations within a sample of 500 PSP's measured in SPBL's 2003/2004 summer inventory. Table 5.1 summarises the main variables.

Table 5.1: Summary of data used for modelling stand leaf area index relationships. Each mean value presented is the mean ± standard error from 141 plots.

Variable	Mean	Standard Deviation	Minimum	Maximum
Understorey (%)	5.82	10.97	0	60
LAI (m^2 / m^2)	2.04	0.35	0.73	2.85
Elevation (m)	219.5	199.27	0	660
Aspect (deg)	144.57	122.63	1	360
Slope (deg)	5.87	6.8	2.5	32.5
Age (years)	21.53	6.26	9.57	30.5
Stocking (stems / ha)	579.43	187.8	150	1650
Tree Mean Top Height (m)	24.14	6.53	10.13	34.76
Stand Basal Area (m^2 / ha)	46.15	20.92	5.19	104.23
Average DBH (cm)	28.19	6.13	13.33	42.97
Maximum DBH (cm)	41.5	10.29	19.5	69.3

To evaluate and test equations, data were prepared in two sets. The first set contained 109 plot measurements for 35 stands. This set was used for fitting canopy LAI equations at a plot level. The second set comprised 32 independent plot measurements for 19 different stands. This second data set was used for validation and testing of final equation forms.

Model Fitting

Equations were fitted using non-linear least-square procedures with SAS software (SAS Institute Inc., 1990). Mean Square Error (MSE) and graphical residual patterns were used as the selection criteria to judge model performance. Plots of residuals against predicted values and against independent variables were examined for bias and residuals were tested for normality using the UNIVARIATE procedure in the SAS statistical package.

To model canopy LAI two non-linear equation forms were examined, the Schumacher (1939) (equation 5.1), and reverse exponential (Mason, pers. comm.) (equation 5.2). The sapwood pipe theory (Waring *et al.*, 1982) suggests that leaf area index should be related to basal area development, and the Schumacher equation is often the best fit to basal area data within even-aged forest stands (Woollons and Whyte, 1990). Effects of mensuration measurements, percent understorey estimates, and climatic variables were considered and formulated into the selected model when significant. A form of the Schumacher (1939) sigmoid yield equation was selected as the general form to model canopy leaf area (Equation 5.1). This model form was chosen as it has been shown to closely follow basal area growth of radiata pine plantations (Zhao, 1999; Snowdon *et al.,* 1999; Wollons *et al.,* 1997), and was hypothesised to closely follow the growth of canopy leaf area in these stands. The model form can be written as,

$$Y = e^{\alpha - \frac{\beta}{X^{\gamma}}} \qquad (5.1)$$

where: Y = dependent variable, X = independent variable, and α, β, and γ = parameters in the equation.

A second form of non-linear equation, a variation on the reverse exponential that actually reaches its asymptote (Mason, pers. com.), was also tested and can be written as,

$$Y = \alpha - (\alpha - \beta)e^{-\gamma(1-x)} - (\alpha - \beta)(1 - X)e^{-\gamma} \qquad (5.2)$$

where: Y = dependent variable, X = independent variable, and α, β, and γ = parameters in the equation. This equation offers complete independence between parameters representing maximum, minimum, and slope.

Model Validation

Validation data comprised 32 independent plot measurements for 19 different stands. The models were examined quantitatively, assessing model behaviour with the validation data set. The procedure involved graphical displays and statistical tests of Average Model Bias (AMB) and Model Efficiency (EF) (Loague and Green, 1991). Potential correlation was detected with inspection of graphical plots of residual versus predictions and

explanatory variables. Model residual errors or the difference between predictions made with Models 1 and 2 and observations of LAI were examined for trends that indicated bias.

Model performance has been described using AMB and EF (Loague and Green, 1991; Zhao, 1999). Average model bias (Equation 3) is an average of errors for all predictions. An AMB of 0 would indicate a model with no bias. Model efficiency (Equation 4) is also a measure of model performance, but a high value of EF (maximum value of 1) indicates a model of perfect fit, an EF value of 0 indicates a model of poor fit where the average value would model the relationship as well, and finally a negative EF result indicates an even poorer fit than the average value.

$$AMB = \frac{1}{n}\Sigma\left(Y_i - \hat{Y}_i\right) \tag{5.3}$$

$$EF = 1 - \frac{\Sigma\left(Y_i - \hat{Y}_i\right)^2}{\Sigma\left(Y_i - \overline{Y}_i\right)^2} \tag{5.4}$$

5.3 Results

Fitted Models

The simple forms of equations 1 and 2 produced MSE values of 0.069 and 0.079 respectively, when used to predict leaf area index with stand age. Two final models of stand leaf area were developed (equations 5.5 and 5.6) using the general form of the Schumacher (1939) sigmoid equation. Equation 5.5 (model 1) was developed to be as simple as possible and rely on only general stand information that can be easily estimated for a site when inventory and physiological data are not available. Model 1 was designed to depend on site variables such as slope, aspect, elevation, stand age, and percent understorey. Equation 5.6 (Model 2) was intended to be used when additional inventory and physiological data were available. This model was intended to include variables such as inventory data stand measurements, and climatic data.

Equation 5.5 was developed using stand age and elevation as independent variables. A final model form is written as,

Model 1 $$LAI = e^{(\alpha_0+\alpha_1 elevation+\frac{\alpha_2 elevation^2}{100000}-\frac{\beta}{age^{\gamma_0+\gamma_1 elevation}})} \quad (5.5)$$

where: *LAI* = canopy leaf area index, elevation of the plot (metres), *age* = stand age (years) α_0, α_1, α_2, β, γ_0, and γ_1 = parameters in the regression equation (Table 5.2).

Table 5.2: Parameter estimates, standard errors, and confidence limits for equation 5.5

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits	
α_0	0.8137	0.0217	0.7707	0.8567
α_1	-0.00087	0.000237	-0.00134	-0.0004
α_2	0.1508	0.0426	0.0663	0.2353
β	751802	2718946	-4640619	6144223
γ_0	6.422	1.5684	3.3114	9.5326
γ_1	-0.00101	0.000298	-0.0016	-0.00042

Residual plots for Model 1 (Equation 5.5) are shown in Figures 5.2 and 5.3. Equation 5.6 was developed using mean top height and elevation as independent variables. A final model form was:

Model 2: $$LAI = e^{(\alpha_0+\alpha_1 elevation+\frac{\alpha_2 elevation^2}{100000}-\frac{\beta}{MTH^{\gamma}})} \quad (5.6)$$

where: *LAI* = canopy leaf area index, elevation of the plot (meters), *MTH* = mean top height (meters) α_0, α_1, α_2, β, and γ = parameters in the regression equation (Table 5.3).

Table 5.3: Parameter estimates, standard errors, and confidence limits for equation 5.6 (Model 2).

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits	
α_0	0.8232	0.0218	0.78	0.8664
α_1	-0.00085	0.000224	-0.00129	-0.00041
α_2	0.1475	0.0389	0.0703	0.2247
β	121604	276619	-426946	670153
γ	4.9897	0.8981	3.2087	6.7707

Residual plots for Model 2 (Equation 5.6) are shown in Figures 5.2, and 5.4. Slope, aspect, basal area, percent understorey, and average rainfall were not significantly correlated with LAI. The final forms of models 1 and 2 produced MSE values of 0.0585 and 0.0533 respectively. General model forms are shown in Figure 5.2.

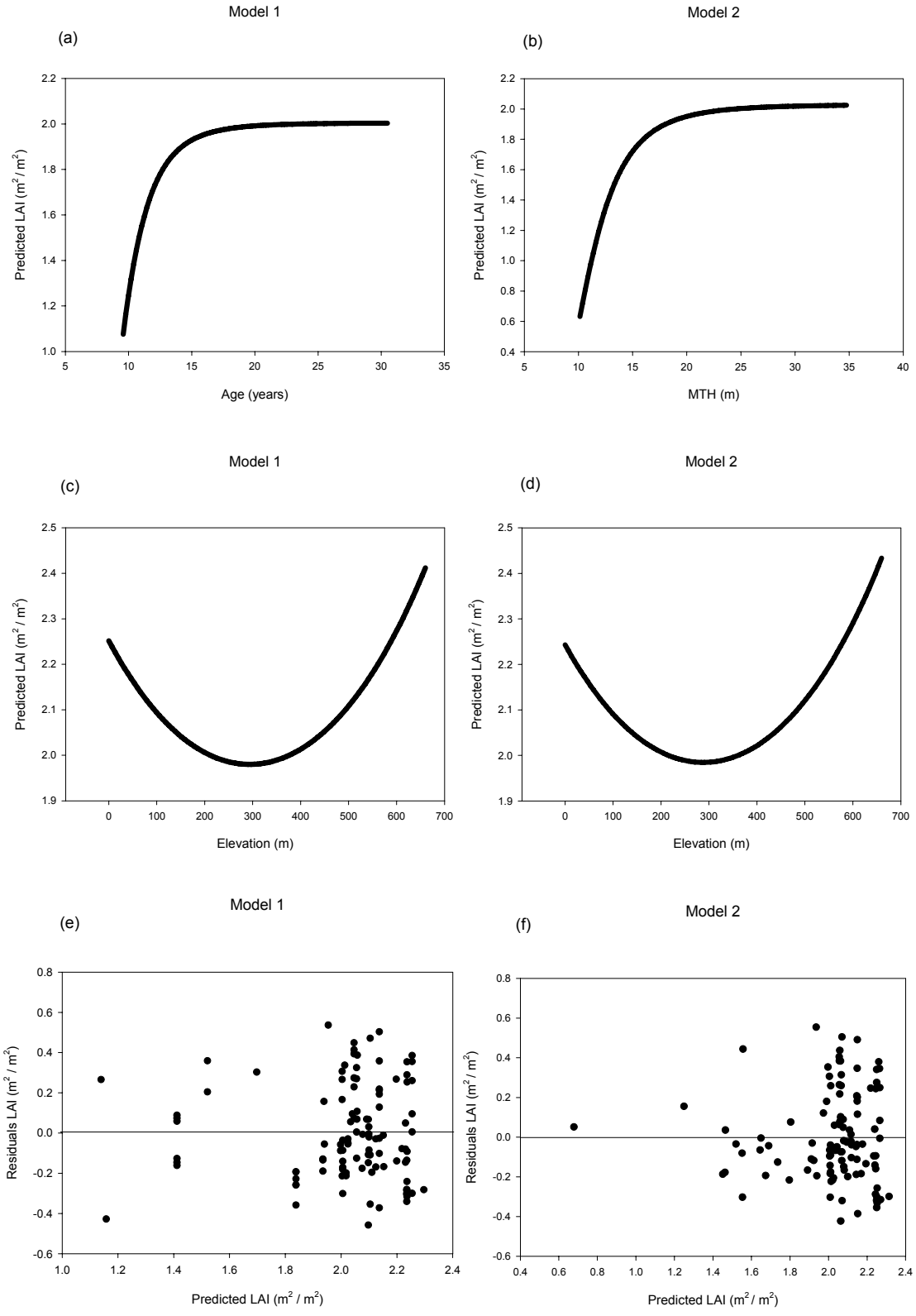


Figure 5.2: General model form and residual plots for models 1 and 2 for (a) predicted LAI from model 1 vs. stand age (holding elevation constant at average = 219.5), (b) predicted LAI from model 2 vs. mean top height (holding elevation constant at average = 219.5), (c) predicted LAI from model 1 vs. elevation (holding stand age constant at

21.53), (d) predicted LAI from model 2 vs. elevation (holding mean top height constant at average = 24.14), (e) and (f) residuals vs. elevation for model 1 and 2.

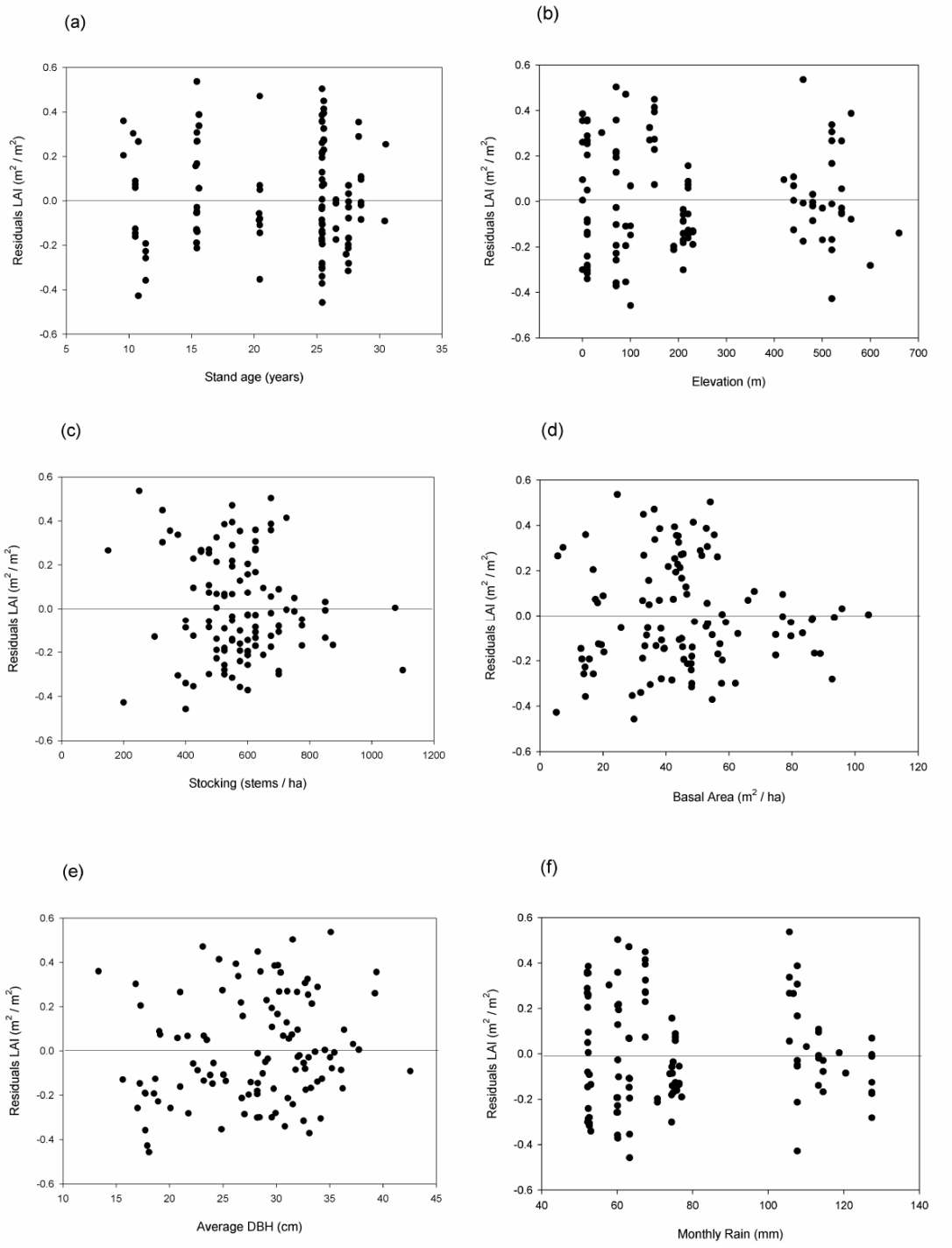


Figure 5.3: Model 1 fitting residual plots for (a) residuals vs. stand age, (b) residuals vs. elevation, (c) residuals vs. stocking, (d) residuals vs. basal area [plot indicates bias], (e) residuals vs. average DBH, and (f) residuals vs. monthly average rain.

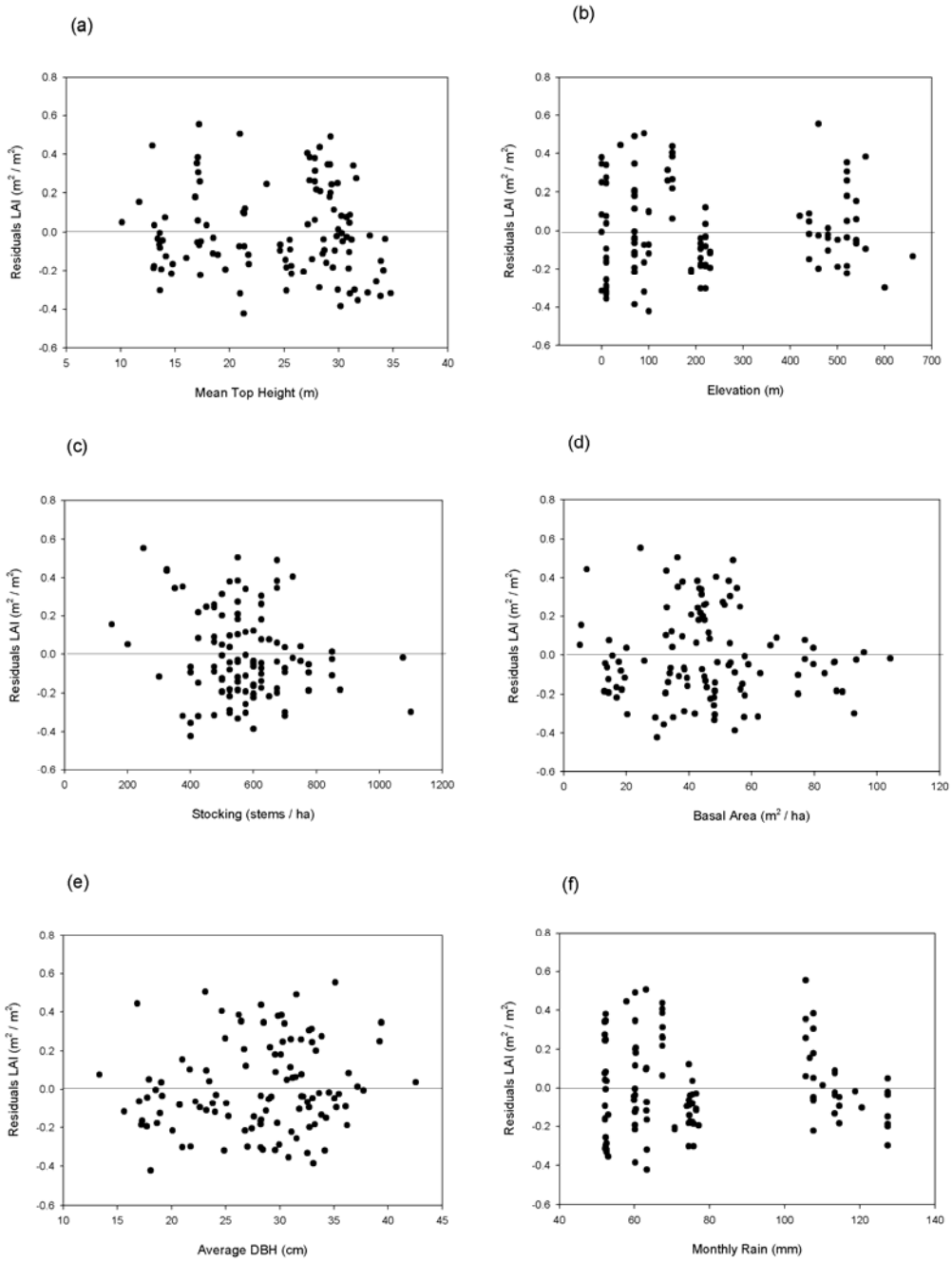


Figure 5.4: Model 2 fitting residual plots for (a) residuals vs. mean top hight, (b) residuals vs. elevation, (c) residuals vs. stocking, (d) residuals vs. basal area [less bias than model 1], (e) residuals vs. average DBH, and (f) residuals vs. monthly average rain.

For both models, model efficiency measures were calculated to validate models against an independent data set. For model 1, EF was calculated at 0.587, and AMB at 0.060, while model 2 had an EF of 0.554 and an AMB of 0.065. Figure 5.5 shows validation residual plots for models 1 and 2

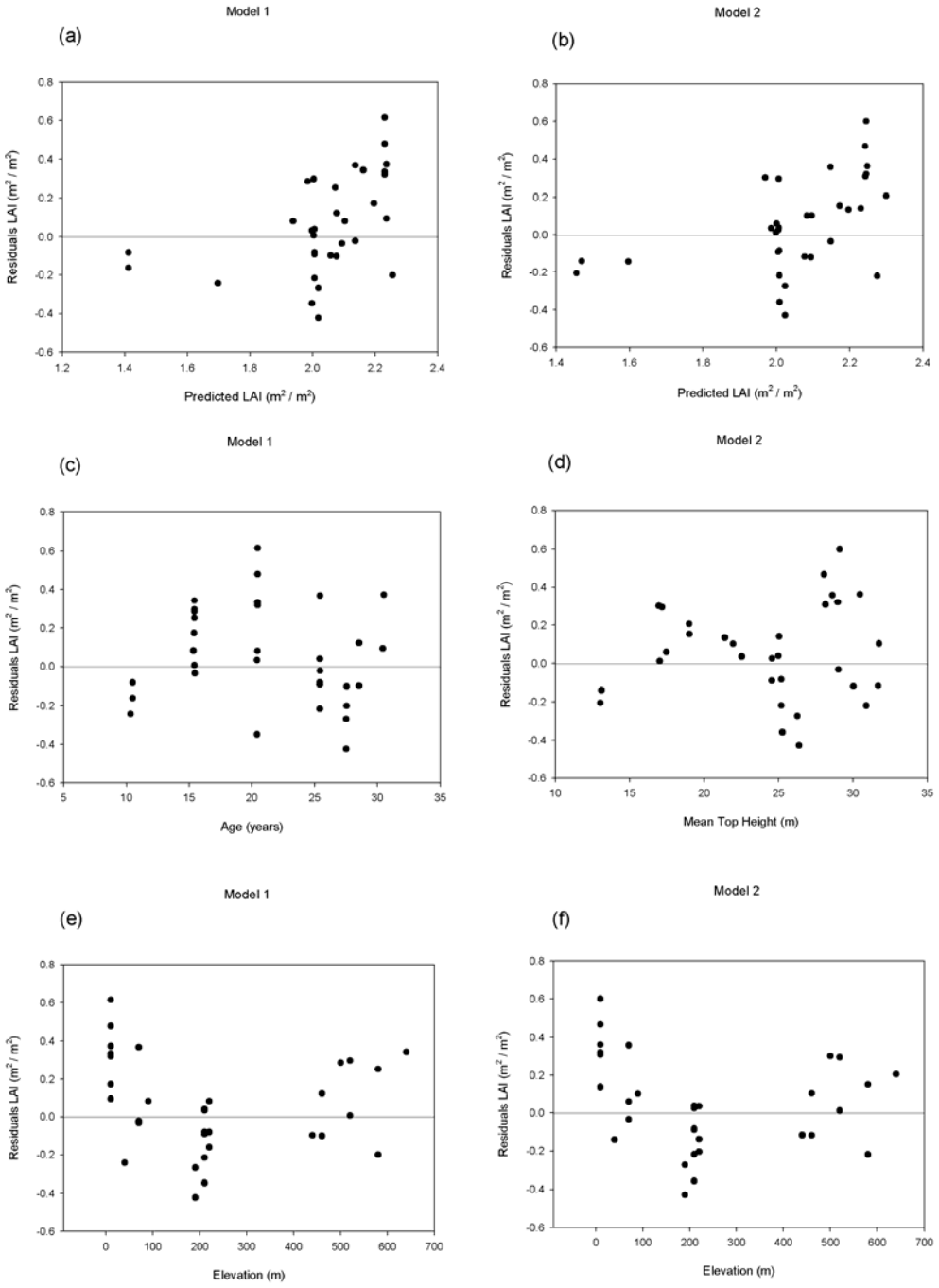


Figure 5.5: Validation plots for models 1 and 2 showing (a) and (b) residuals vs. predicted for models 1 and 2, (b) and (c) residuals vs. age (model 1) [plot indicates bias], and residuals vs. mean top height (model 2) [less bias than model 1], (e) and (f) residuals vs. elevation for models 1 and 2 [both plots indicate bias].

5.4 Discussion

Two final models were developed to predict stand-level LAI from stand state variables. LAI estimated with these models can be used as an input in ecosystem models that simulate carbon balance and hydrological cycles. The two final models were surprisingly simple in form, depending only on stand age, mean top height, and elevation. Results from this study demonstrate the insensitivity of LAI to stocking in older stands. During model fitting, there were no apparent patterns in residual plots against other possible predictor variables, such as stocking and rainfall (Figures 5.3(c and f), and 5.4(c and f). This pattern of simplicity has also been reported by Pinkard and Neilsen (2003), who found similar results for stand leaf areas in plantation-grown eucalypts. They reported canopy leaf areas that were nearly independent of plantation spacing, having only a moderate effect at spacings below 833 trees/ha. These results are also supported by Savill *et al.* (1997) who summarised the general effects of spacing in even aged stands as: once canopy closure has occurred and a site is being fully used in widely spaced stands, dry matter increment is similar to that in close spacings on the same site.

Validation of the two final models with an independent data set revealed patterns in the residual plots that were not apparent in the model building process. This potential bias in residuals may have been due to the small number of validation data points used. There were general patterns in the residuals plots against stand age (Figure 5.5(b)), and mean top height (Figure 5(d)). The introduction of mean top height, which was used in model 2 rather than stand age, used in model 1, improved bias in residual plots. There was also a general increase in residuals with decreasing elevation (Figures 5.5(e) and 5.5(f)). Canopy leaf area was found to increase on some sites at lower elevations. This effect can be viewed by increasing residual errors for both models 1 and 2 at elevations very close to zero, in coastal forests (Figures 5.5(e) and 5.5(f)). This increased growth may have

been due to tree roots tapping into ground water sources located very close to the surface. The general effect of underground water sources distances decreasing with decreasing elevation can be seen in Figure 5.6, which shows depth to ground water in wells in the Canterbury region (Environment Canterbury, 2004). Future research in the areas of canopy leaf area index modelling or growth modelling could focus on including this effect.

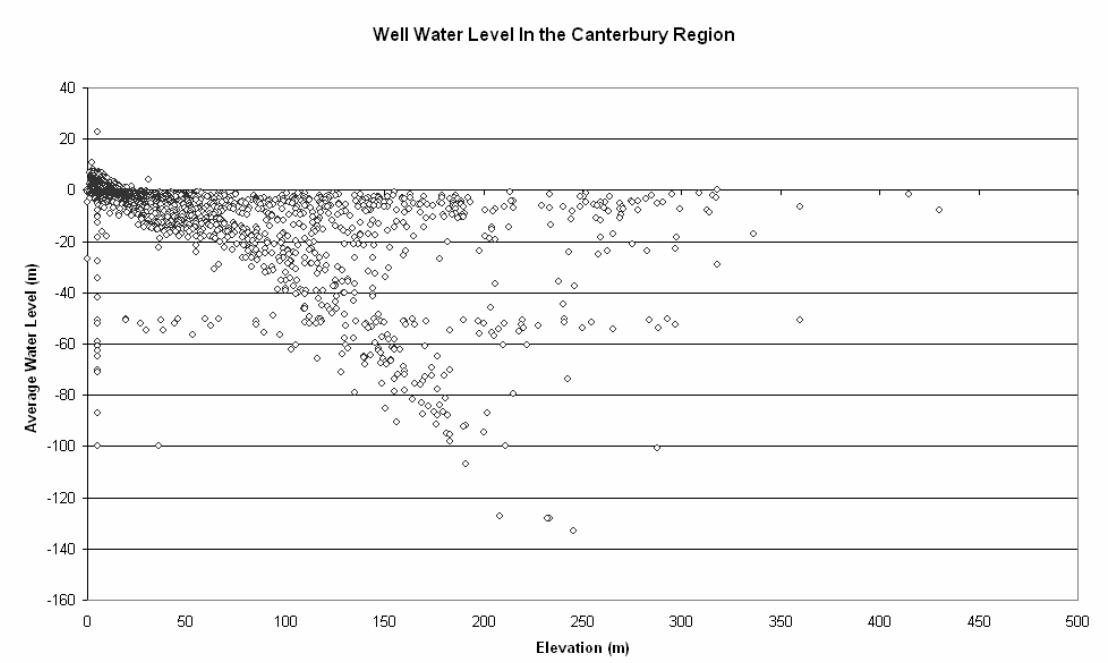


Figure 5.6: Ground water well depth (in the Canterbury region) plotted against elevation. (Environment Canterbury, 2004). Positive well depths indicate a positive well pressure.

The developed models are limited to plantation-grown radiata pine. The models should be used in geographic regions that have similar climatic and soil characteristics as those within the study area, and should only be used for simulation in stands that range from 10 to 32 years of age. There are discrepancies between actual LAI values, and those estimated with hemispherical photography usually due to canopy clumping and non-differentiation of foliage and non-foliage elements in the canopy (Martens *et al.,* 1993). Any use of models 1 and 2 that requires an absolute measure of leaf area will also require a calibration of model results with actual values of LAI measured with destructive sampling. Mason *et al.* (in prep) have calibrated both hemispherical photographs and LAI2000 (LI-COR, 1991) measurements of LAI with destructively sampled estimates of LAI for radiata pine growing in Canterbury. To convert LAI from models 1 and 2 to

actual LAI (one sided) estimated from destructive sampling Mason *et al.* (in prep) suggest a linear adjustment described by equation 5.7:

$$LAI_{ds} = LAI \cdot 3.676737832 - 1.386871774 \qquad (5.7)$$

Where LAI_{ds} = LAI obtained from destructive sampling, and LAI = LAI obtained from models 1 and 2.

5.5 Conclusions

Two models of LAI in radiata pine plantations in the Canterbury region were developed with the equations:

$$LAI = e^{(0.8137 - 0.0008 * elevation + \frac{0.1508 * elevation^2}{100000} - \frac{751802}{age^{6.422 - 0.00101 * elevation}})}$$

Model 1

$$LAI = e^{(0.8232 + -0.00085 * elevation + \frac{0.1475 * elevation^2}{100000} - \frac{121604}{MTH^{4.9897}})}$$

Model 2

Independent variables included elevation, stand age, and mean top height. The following variables were tested and were not significantly related to LAI: stocking, basal area, average diameter at breast height, and average annual rainfall. Residuals of model 1 ranged between -0.46 and 0.53, with an average residual of -0.0016 (MSE = 0.0585), while residuals of model 2 ranged between -0.42 and 0.55, with an average residual of -0.0005 (MSE = 0.0533). Testing with an independent dataset indicated that the models were more biased than expected against elevation, but overall performance of the model was similar to that found during estimation of model coefficients. Model 2 is

recommended for use as it displayed a reduced overall bias and a reduction in bias with respect to independent variables when tested with an independent data set.

CHAPTER 6

Incorporating an Index of Root Zone Water Balance into the Existing Growth and Yield Model CanSPBL(1.2)

6.1 Introduction

Water balance is a well established indicator of growing potential and has been found to be a main constraint for tree development on dry forested sites. Reduction in the availability of water decreases stem growth by restricting physiological processes such as leaf area development, photosynthesis, and stomatal conductance (Boomsma and Hunter, 1990). An index for water stress (the water stress integral) has been shown by Myers (1998) to account for almost all of the variation in basal area increment of *Pinus radiata* subjected to a range of irrigation and fertilization treatments. Given the importance of water availability in regulating productivity and basal area growth on dry sites, a model of root zone water balance may provide a useful process-based approach that can be incorporated into a classical growth and yield model.

Incorporating a model of root zone water balance into a classical growth and yield model may increase accuracy of tree growth prediction, and make the model more robust under changing climatic conditions. This chapter will discuss a root zone water balance model that was used to estimate water balance for all forests in the Selwyn Estate. An index for available water at the each site will be used as a variable to predict future forest growth. This index was incorporated into the equations that comprise the growth model canSPBL(1.2). This chapter will first examine the modelling strategy used to predict root zone water balance, discuss the input data for the model, parameterization of the model, validation of the model used to predict root zone water balance, and then the results of incorporating that index into canSPBL(1.2).

6.2 Water Balance Modelling

3PG Water Balance Model
Adapted from (Landsberg and Waring 1997) and (Sands and Landsberg 2002)

One of the water balance models considered to calculate an index of water use over the measurement interval is the model used in 3-PG (Landsberg and Waring, 1997). This model is a single layer soil-water-balance model that operates on a monthly time step. Monthly rainfall (plus irrigation) is balanced against monthly evapotranspiration computed using the Penman-Monteith equation (Landsberg and Gower, 1997, p. 76.) The general equation can be written as;

$$\theta_T = \theta_{T-1} + P - I - E - D \qquad (6.1)$$

Where: θ_T is the root zone water balance at time T (mm), θ_{T-1} is the root zone water balance at time (T-1) in the previous month (mm), P is precipitation (mm), I is the canopy interception (mm), E is evapotranspiration (calculated using the Penman-Monteith equation,) and D is the drainage of water from the soil (mm).

The model is initialized with soil water content = maximum available water (Θ mm) in the rooting zone. This is dependent on the water holding characteristics of the soil and the rooting depth of the trees (see Landsberg and Gower, 1997). Available water is set to appropriate starting values for each site. The moisture ratio (r_θ) for the stand is calculated as:

$$r_\theta = \frac{\text{Current soil water content + water balance}}{\text{Available Water}} \qquad (6.2)$$

The moisture ratio r_θ is used in the calculation of a soil water modifier (f_{sw}), which is used to modify estimates of canopy conductance (g_c) that will be discussed later.
The water balance in any month will be reduced if transpiration exceeds precipitation, and vice versa. If the numerator of the expression for r_θ exceeds Θ, the excess water is assumed to have run off or drained out of the system. If it is negative, $r_\theta = 0$.

Canopy interception is a fixed percentage of rainfall when the canopy exceeds a threshold leaf area index (LAI) in the 3-PG model. When P > 2mm, and LAI ≥ 3, then I = 0.15*P. If P< 2 mm, then I = P. And if P > 2mm and 0 < LAI < 3, then I = 0.05*LAI*P.

Vapour pressure deficit, available soil water and stand age are assumed to affect stomatal conductance. Canopy conductance (g_c (m s^{-1})) is determined from a nominal stomatal conductance scaled by f_{age} and by the lesser of the environmental modifiers f_{sw} and $f_{VPD,}$ and increases with increasing canopy LAI up to a maximum canopy conductance ($g_{c_{\max}}$ (m s^{-1})).

The equation for canopy conductance is given by:

$$g_c = g_{c_{\max}} \cdot e^{(-k \cdot VPD)} \tag{6.3}$$

Where k is a factor based on the relationship between stomatal conductance and vapour pressure deficit = 2.5 (value taken from Landsberg and Waring 1999) and *VPD* is the saturation vapour pressure (kPa).

The age modifier , f_{age} is calculated from:

$$f_{age} = \frac{1}{1 + \left(F_a \Big/ 0.95 \right)^{n_{age}}} \tag{6.4}$$

Where F_a is the relative age of a forest or plantation denoted by the ratio of actual age (in years) to the maximum age likely to be attained, n_{age} is a parameter to control the rate of change of the function and the default value from the 3-PG model is $n_{age} = 4$.

The vapour pressure deficit modifier , f_{VPD} is calculated from:

$$f_{VPD} = e^{(-k \cdot VPD)} \tag{6.5}$$

Where k is a factor based on the relationship between stomatal conductance and vapour pressure deficit = 2.5 (value taken from Landsberg and Waring 1999) and *VPD* is the saturation vapour pressure (kPa).

The soil water modifier, f_{sw} is calculated from:

$$f_{sw} = \frac{1}{1 + \left[\left(1 - r_\theta\right) / c_\theta \right]^{n_\theta}} \qquad (6.6)$$

Where c_θ and the power n_θ take different values for different soil types. Landsberg and Waring (1997) suggest c_θ = 0.7, 0.6, 0.5, and 0.4 for sand, sandy-loam, clay-loam, and clay soils respectively, and n_θ = 9, 7, 5, and 3 for the same soil types.

Watt Model

A second model was also considered in this study to estimate root zone water balance. It has been used previously by (Whitehead *et al.*, 2001; Watt *et al.,* 2003) and is formulated as follows:

$$W_i = W_{i-1} + P_i - E_{ti} - E_{twi} - E_{gi} - F_i \qquad (6.7)$$

where P_i is rainfall, E_{ti} the transpiration from the tree canopy, E_{twi} the evaporation from the wet tree canopy, E_{gi} the evaporation from the soil surface, and F_i drainage from the root zone, and surface runoff was assumed to be insignificant (Whitehead *et al*., 2001).

Some tree-specific parameters for the water balance models were used: boundary layer conductance for trees (*gBT*) = 0.2 (m / sec), the intercept of net radiation for trees (*qaT*) = -90 (watts / m^2), and the slope of net radiation for trees (*qbT*) = 0.8.

6.3 Water Balance Validation

The 3PG, and Watt water balance models were tested against a validation dataset that contained measurements taken within the study area. The validation dataset came from a study done by Arneth *et al.,* (1998), which contained 49 measurements of volumetric root-zone water content taken at Balmoral Forest, located 100-km Northwest of Christchurch New Zealand, (42°52'S, 172°45'E at an elevation 198 m above sea level) between 1994 to 1996. A residual analysis of both models was run across the validation dataset. This residual analysis was conducted on a cumulated deficit basis over a given measurement interval. This was done to mirror the models final use in a growth and yield modelling context where the accumulated water deficit over a measurement interval is expected to describe variations in stand growth and diameter distribution. Table 6.0 shows the actual and estimated accumulated water deficit for both the 3PG and Watt water balance models. An analysis of mean square error (MSE Eq. 6.2) was used as the final selecting criterion to determine the water balance model used in further analysis.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y - Y'\right)^2 \qquad (6.2)$$

Where Y is the measured accumulated root zone water deficit (mm), and Y′ is the modelled accumulated root zone water deficit.

Table 6.0: Measured and estimated accumulated water deficit (mm) from 3PG and Watt water balance models.

Interval (month/year)	Measured accumulated water deficit (mm)	3PG estimate (mm)	Watt estimate (mm)
10/1987 - 5/1988	273.51	273.57	271.77
6/1988 - 5/1989	417.11	379.52	354.77
6/1989 - 3/1990	164.72	128.79	144.34

The 3PG water balance model produced a MSE of 903.1, while the Watt model produced a MSE of 1434.96. While it is recognized that this validation was limited and not

adequate to make a clear distinction, the 3PG model was selected as the final water balance model for further analysis based on the MSE calculated within this validation.

6.4 Climatic Inputs for Water Balance Modelling

Climatic inputs to the water balance model for each site include minimum temperature, average temperature, vapour pressure deficit, precipitation, and solar radiation. These inputs are based on a system of corrected estimates from BIOCLIM (Leathwick *et al.* 1998) climate surfaces. BIOCLIM is a set of surface equations for New Zealand that estimate climatic variables on a ten year average based on interpolation between climate station measurements (Leathwick *et al.* 1998). BIOCLIM outputs monthly averages of temperature, wind, rainfall, radiation, and average humidity extremes across New Zealand based on location. To simulate the monthly variation in actual climate readings, as opposed to mean monthly values over a 10-year period, the long-term mean monthly climate values derived from BIOCLIM were rescaled using monthly climate measurements from a nearby weather station. These estimates were corrected by reference point measurements made at Christchurch Airport (S43.5°, E172.55°, elevation 37m). Differences between climatic estimates from BIOCLIM and measurements at Christchurch Airport were calculated for each month included in the study from 1984 – 2004. Corrections were added to long term monthly estimates at each permanent sample plot. So instead of using the same monthly averages every year, a new value was calculated every month based on actual climate at Christchurch Airport to reflect years of extreme climatic events such as droughts. Similar approaches of correcting long term averages with local measurements using a ratio adjustment have been used by Tickle *et al.* (2001), and Snowdon *et al.* (1999).

The approach of adjusting long term climate estimates from BIOCLIM with local measurements via a difference method as opposed to a ratio method was validated with an independent data set of climate measurements at Darfield weather station (E172.13° S43.48°).

The difference method of climate adjustment is written as:

$$B_{2adj} = B_2 - \left(B_1 - m_1\right) \qquad (6.3)$$

The ratio method of climate adjustment is written as:

$$B_{2adj} = \frac{B_2 \cdot m_1}{B_1} \tag{6.4}$$

Where B_{2adj} is the adjusted BIOCLIM estimated for a given site, B_2 is the BIOCLIM estimate unadjusted for the same site, m_1 is the climatic measurement at Christchurch airport, and B_1 is the BIOCLIM estimate for the Christchurch airport.

The validation across all BIOCLIM variables for monthly averages of, maximum temperature, minimum temperature, average temperature, solar radiation, vapour pressure deficit, and wind speed, showed less bias overall with the difference method of adjustment (Table 6.0.1, Figures 6.0-6.1).

Table 6.0.1: Residual mean square errors for the ratio and difference methods of climate estimate adjustment.

Climate Variable	**Ratio method MSE**	**Difference method MSE**
Rain (mm)	458	402
Average Temperature (°C)	2.32	2.33
Maximum Temperature (°C)	2.73	2.69
Minimum Temperature (°C)	0.71	0.78
Solar Radiation (MJ /m^2)	0.908	0.907
Vapour Pressure (kPa)	0.002	0.001
Wind Speed (km / hr)	0.21	0.22

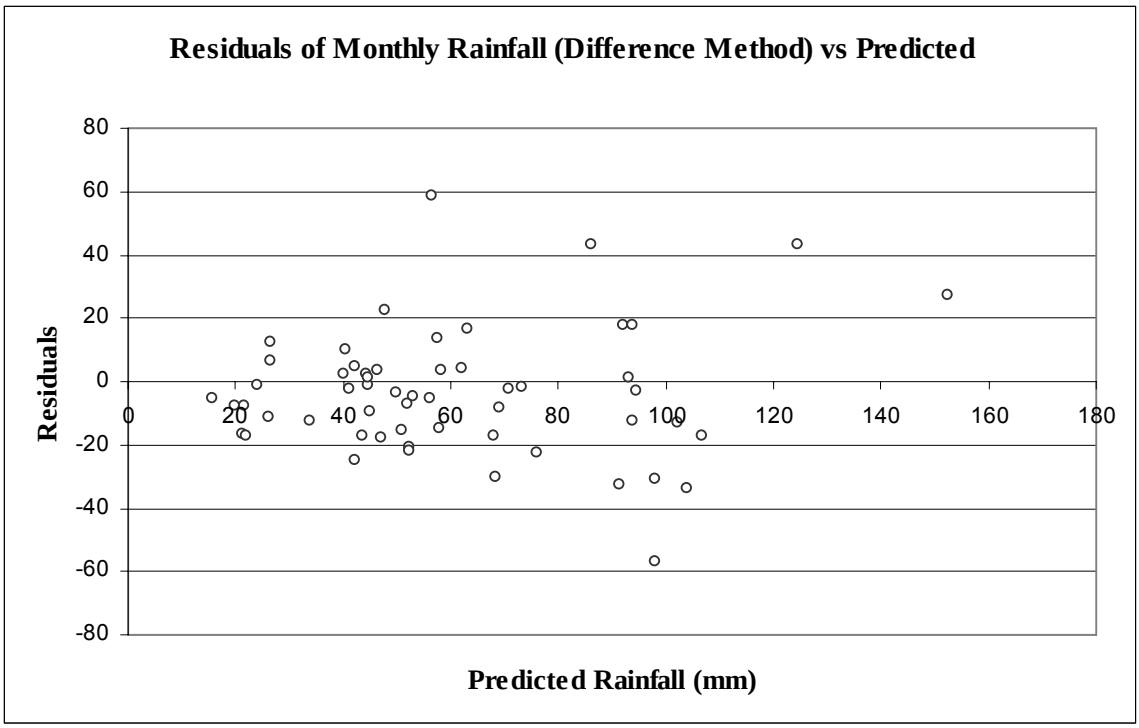


Figure 6.0: Residuals of monthly rainfall vs. predicted.

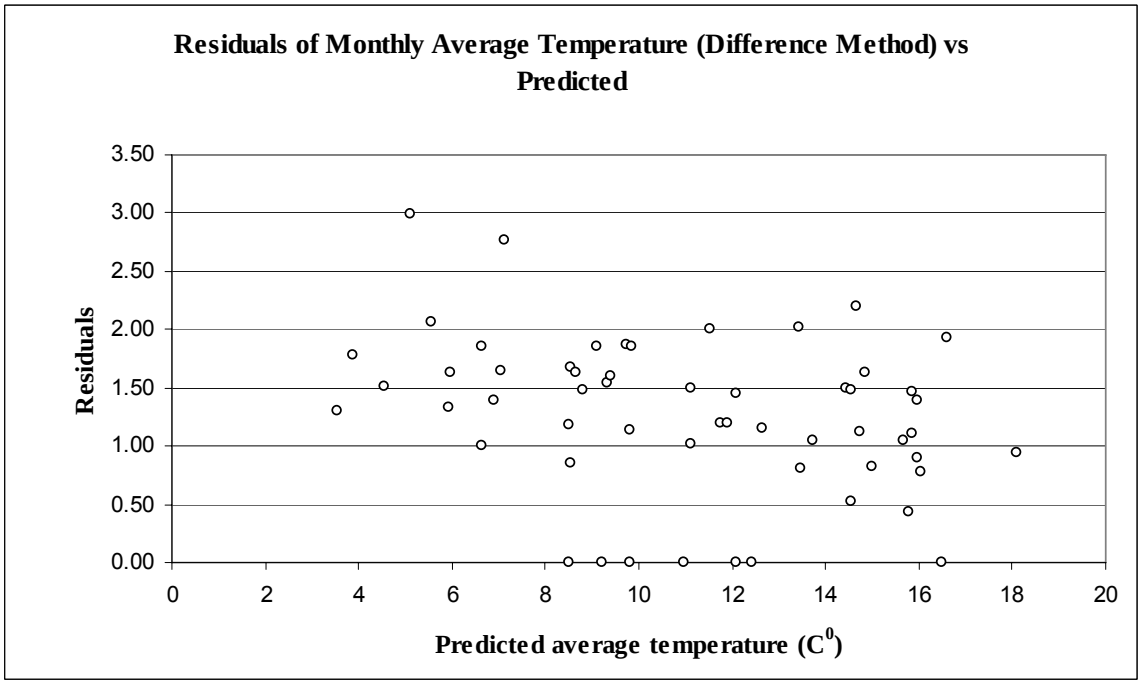


Figure 6.1: Residuals of average monthly temperature vs. predicted.

Rainfall estimates obtained by either the difference or ratio method of correction showed a high bias (Table 6.0.1). A comparison of BIOCLIM estimates and actual measurements at twelve stations in the study area showed a bias in over prediction of rainfall at all

stations (Figure 6.2). Twelve monthly average rainfall regressions of BIOCLIM values versus monthly averages (Table 6.2) for 27 weather stations around the study area were conducted to locally correct for this bias. Reference datasets were obtained from the National Institute of Water and Atmospheric Research (Table 6.1).

Table 6.1: Weather stations used in monthly rainfall regressions.

Station	Average Annual Rain	Measurement Start	Measurement Finish	Elevation (m)
Shirley	570	1967	2002	6
Windsor	582	1978	1997	5
Waimakariri	587	1967	1999	5
Christchurch, Bromley	591	1961	1990	9
Woodend	609	1967	2002	6
Christchurch	611	1945	2003	37
Greenpark	624	1956	2004	2
Papura	636	1925	1988	50
Prebbleton	641	1969	2004	21
Burnham	642	1956	2004	63
Lincoln-Broadfield	646	1972	2000	12
Rangiora	666	1965	1998	23
Dunsandel	687	1973	1998	73
Te Pirata	716	1937	2004	144
Somerton	740	1950	2004	137
Eyrewell	752	1942	1989	158
Darfield	782	1939	2002	195
Hororata	850	1960	2004	213
Homebush	867	1925	2002	244
Hororata West	909	1967	2002	287
Highbank	932	1953	2002	230
Windwhistle, Rockwood	1073	1967	1973	366
Woodlands	1083	1967	1991	366
Glendore	1100	1969	1976	360
Dalethorpe	1109	1967	1980	396
Snowdon	1125	1999	2004	560
High	1150	1967	2002	457

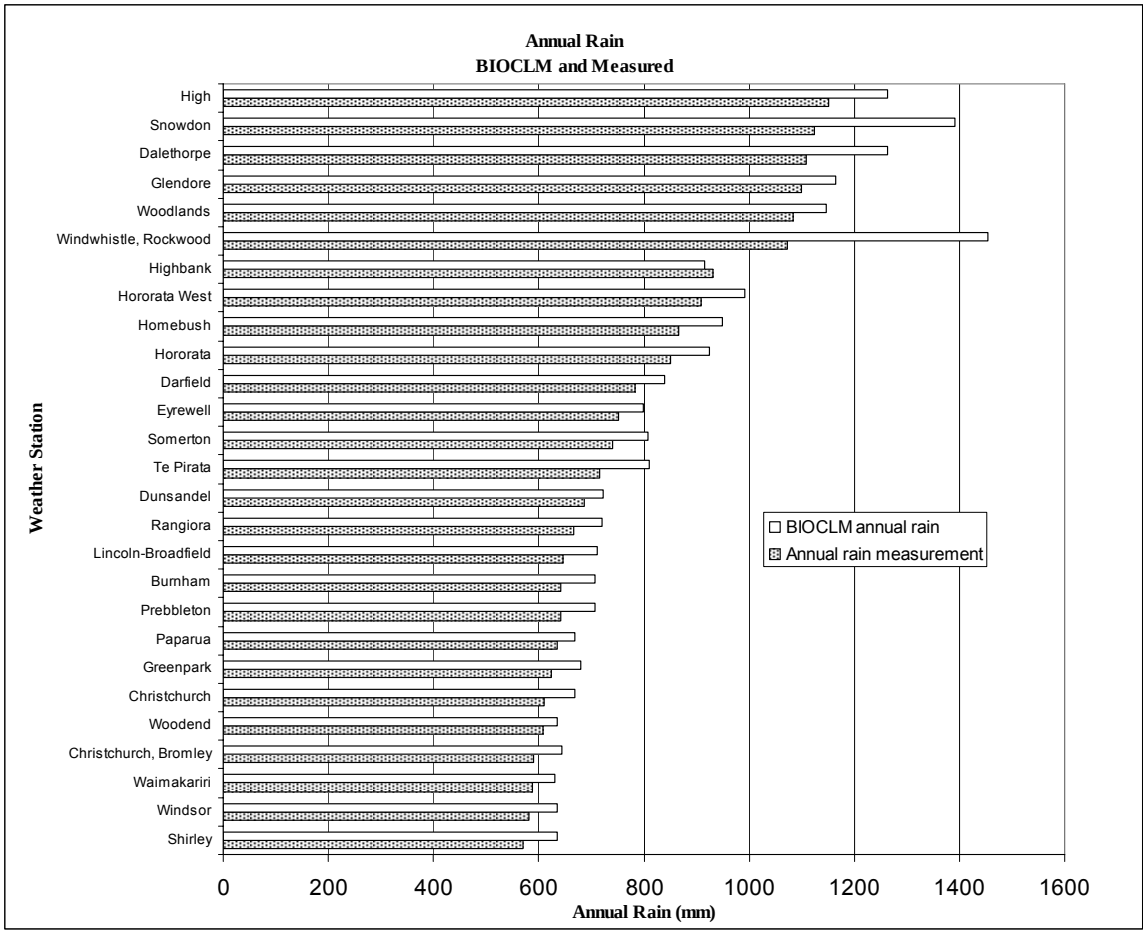


Figure 6.2: A comparison of BIOCLIM estimates and actual measurements at twenty seven stations in the study area.

Table 6.2: Monthly rainfall regression coefficients and R^2 , where X = average monthly rain estimated from BIOCLIM, and Y = adjusted average monthly rainfall (mm).

Month	Intercept	Slope	R^2
1	10.049	0.9947	0.95
2	3.7301	0.9958	0.92
3	-2.1475	1.2091	0.79
4	-0.9145	1.2598	0.85
5	19.664	0.9254	0.88
6	-4.3858	1.0402	0.7
7	10.064	0.8808	0.69
8	-3.2399	1.0486	0.74
9	1.256	0.9779	0.86
10	-14.693	1.3208	0.91
11	-0.7513	1.106	0.93
12	-10.17	1.393	0.73

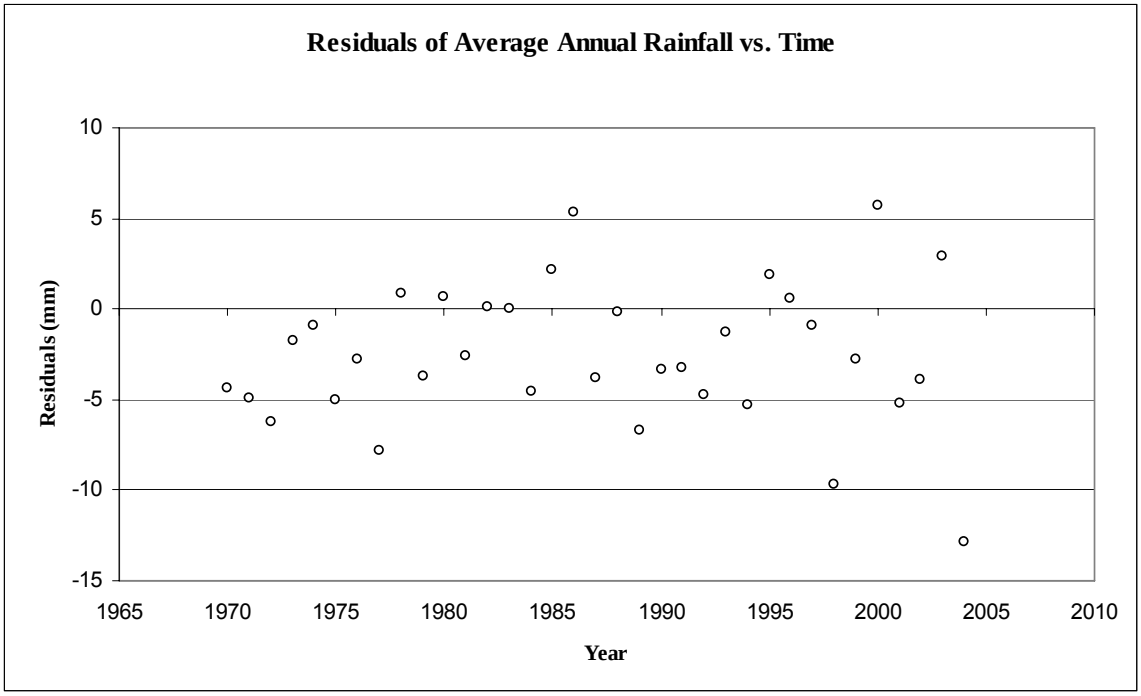


Figure 6.3: Residuals of monthly rainfall, averaged annually vs. time. Points represent monthly rainfall (averaged annually) for 27 stations included within the study area, corrected by monthly rainfall regressions.

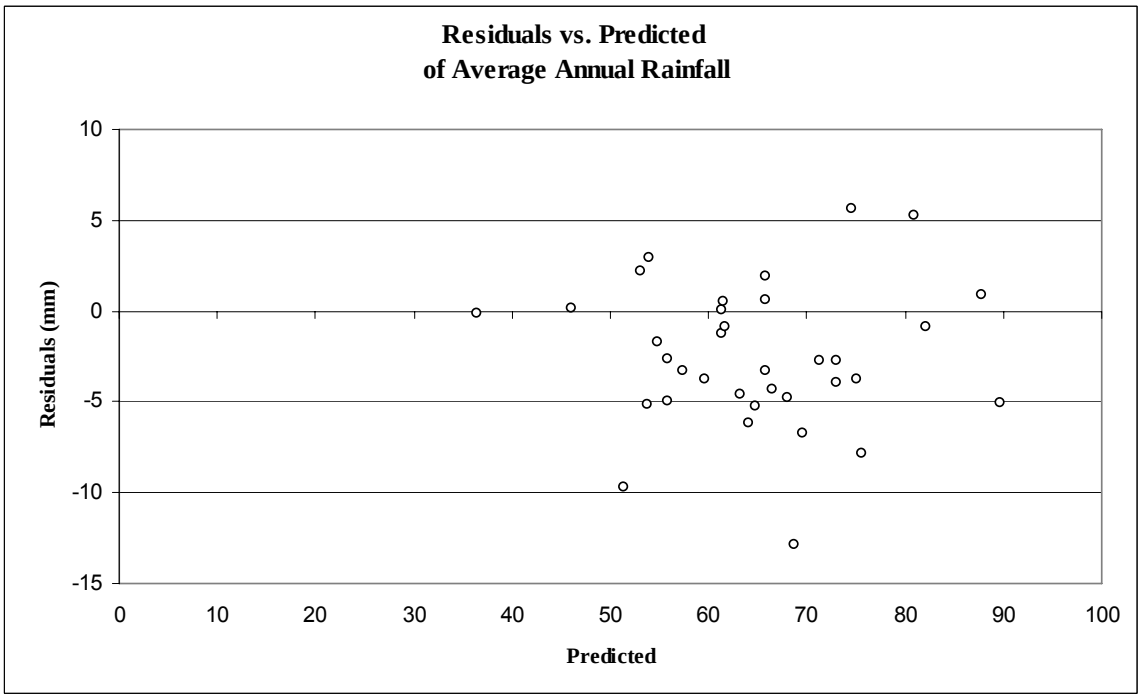


Figure 6.4: Residuals of average monthly rainfall (averaged annually) vs. predicted. Points represent monthly rainfall (averaged annually) for 27 stations included within the study area, corrected by monthly rainfall regressions.

After adjusting rainfall estimates with monthly rainfall regressions listed in Table 6.2, bias and total error for estimated rainfall at 27 met stations was reduced. Estimated monthly rainfall (averaged annually) was within 5mm for most stations (Figure 6.3, and Figure 6.4).

The water balance model will also require input information on soil type, maximum available soil water, and minimum available soil water. Available soil water (ASW) levels and soil classification data were extracted from a local database of New Zealand soil types (Barringer *et al.* 1998) where soil attributes are linked to latitude and longitude.

6.5 Growth Modelling Methods: Incorporating an Index of Water Balance into CanSPBL(1.2)

6.5.1 Methods

Equations used to model growth in CanSPBL(1.2) were updated to incorporate an index of root zone water balance. The water balance index used described the average available soil water deficit over the simulation period, as calculated with the model for water balance used in 3PG (Landsberg and Waring 1997). The water balance model was represented in the Java programming language by Dr E.G. Mason, and he accessed weather data for each plot within a Microsoft Access database table and arranged for output to be placed alongside growth and yield data in a separate Access table. Table 6.3 shows the descriptive statistics of average monthly available soil water deficit (ASW) values for the model fitting data set.

Table 6.3: Descriptive statistics of the average ASW deficit in modelling data set.

3PG Average Water Deficit	
Mean (mm)	32.21
Minimum (mm)	0.00
Maximum (mm)	142.38
Standard Deviation (mm)	24.95
Count	3660

The method used for incorporating the effect of average ASW deficit into the equations used in CanSBPL(1.2) was to plot the residuals of the models against ASW deficit and observe the pattern of residual bias. This bias was then accounted for by adding curvilinear forms of ASW deficit to the models. An example of this is shown in Figure 6.4 where the bias of mean top height residuals plotted against ASW deficit can be corrected in part by a curvilinear model including a form of ASW deficit and ASW deficit squared.

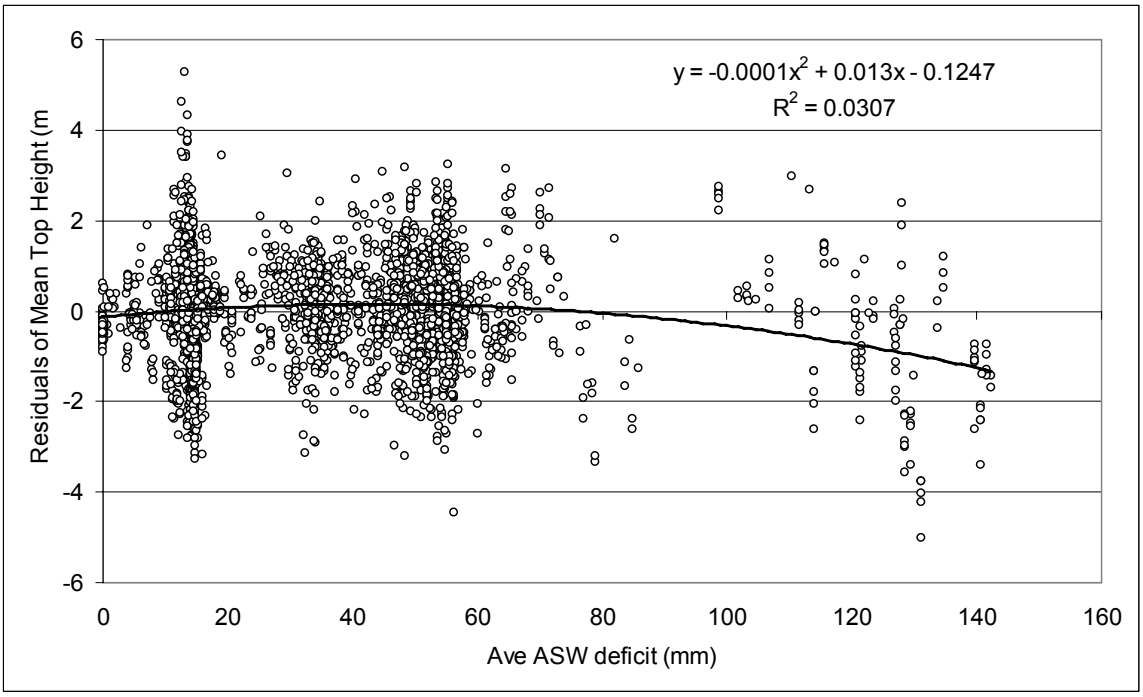


Figure 6.4: Residuals of mean top height vs. average 3PG available soil water deficit.

Residuals of all of the models in CanSPBL(1.2) were first plotted against average ASW deficit. All of these plots indicated the bias could be corrected for by adding ASW deficit and ASW deficit squared to each model. The 4 best initial fit equations where the effects of elevation were initially tested in chapter 4 were fitted again with ASW deficit and ASW deficit squared added to them. These final models were then tested further to see if different forms of elevation and ASW deficit could be added to them in a way that would minimize error and simplify the final model.

To fit equations, methods used by Zhao (1999), of non linear least-square procedures with SAS software (SAS Institute Inc., 2001) were employed. The Mean Square Error

(MSE) and graphical residual patterns were used as the selection criteria to judge model performance. Skewness and kurtosis were often checked for final models to determine the magnitude of residual distributions of normality. Plots of residuals versus predictions and all possible explanatory variables were inspected to check for trends but only the five most important graphs are displayed in this study for each selected equation. The five graphs are residuals versus prediction, age, elevation, time increment, and ASW deficit.

No statistical tests are given in this analysis because repeated measurements have been taken from the basic experimental units (PSP). The consequences of this and a rationale for using auto correlated data in growth and yield modelling are described in chapter 4. Some statistically valid tests are provided in this analysis as a check of the final results by preparing an auto correlation free data set. An auto correlation free data set was used to test the significance of explanatory variables being different from zero for all models and check the normality of residuals with the Kolmogorov-Smirnov test. The Kolmogorov-Smirnov test was chosen to test normality because of the large size of the data set.

To validate the newly established model and test its performance at a plot level, a validation data set was prepared. The main model components of MTH, basal area per hectare, stems per hectare, maximum diameter, and standard deviation of diameter were examined using data at a plot level. One source of data was selected consisting of independent plot measurements chosen randomly before model fitting. The average model bias (AMB), efficiency factor (EF), skewness and kurtosis were calculated for each model component. Graphs of residual patterns were examined to detect bias. Normality of residual plots were also tested using the Kolmogorov-Smirnov test.

5.5.2 Results

Basal Area

The four models with the best initial fit for basal area in chapter 4 were; Polymorphic Gompertz II, Polymorphic Schumacher II, Polymorphic Weibull I, and Polymorphic Weibull II as shown in Table 4.32. The effects of water balance deficit and elevation was added into these models by altering the asymptotic parameter for each equation. These equations were initially fitted with the form of elevation that was used for the final model

of basal area in CanSPBL(1.2) and the effects of water deficit added in as a linear and curvilinear term. The results of initial model fits in terms of MSE are shown in Table 6.4.

The model with the lowest MSE after the addition of the effect of average water deficit and elevation was altered further to see if different correction forms for these effects would improve model fits. These included a linear correction for elevation for PSP's above 250 m, as was used by Zhao (1999), elevation, the square of elevation, a probit correction for the effect of elevation, ASW deficit, ASW deficit squared, and the interaction of elevation and ASW deficit.

Table 6.4: Basal area model fitting results, after the effect of average water deficit and elevation are included in the asymptotic parameter.

Equation with elevation and average 3PG water deficit in asymptote	Basal Area MSE
Polymorphic Weibull I	failed to converge
Polymorphic Gompertz II	16.49
Polymorphic Weibull II	failed to converge
Polymorphic Schumacher II	16.81

The best model fit was found by adjusting the asymptotic parameter of the Polymorphic Gompertz II equation using elevation squared, a dummy variable for elevations above 450m, mean ASW deficit, and mean ASW deficit squared. The final model for basal area has a MSE of 16.53, which is slightly higher than the initial fitting of the Polymorphic Gompertz II model listed in table 6.4. The reason for this result is that this final model does not have the term of elevation alone in the asymptotic parameter as it was insignificant. The final model was fitted without this term and is written as:

$$G_2 = e^{\left(\ln(G_1)e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}+\frac{\left(\alpha_0+\alpha_1\cdot elev^2+\alpha_2\cdot((elev-450)\cdot X)+\alpha_3\cdot ASW+\alpha_4\cdot ASW^2\right)}{10000}\left(1-e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right)\right)} \quad (6.5)$$

Where G_2 is the future basal area (m^2 / ha), G_1 is the initial basal area (m^2 / ha), T_1 is the initial stand age (years), T_2 is the final stand age (years), *elev* is the stand elevation (m), X is a binary indicator variable, $X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450, ASW is the average water deficit over the simulation period calculated with the 3PG water balance model (mm), and β , γ, α_0, α_1, α_2, α_3, and α_4 are parameters whose values are listed in Table 6.5.

The parameter estimates of the final model of basal area are listed in table 6.5. The model showed no signs of bias for basal area / ha when plotted against prediction, age, elevation, and time increment (Figure 6.5).

Table 6.5: Parameters for basal area model (Equation 6.5), standard errors, and approximate 95% confidence limits calculated with 3659 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.153	0.00219	0.1487	0.1573
γ	0.00249	0.000038	0.00242	0.00257
α_0	44311.5	229	43862.6	44760.4
α_1	0.0358	0.000708	0.0344	0.0372
α_2	-31.7278	2.1729	-35.9881	-27.4675
α_3	35.2148	4.363	26.6604	43.7691
α_4	-0.3759	0.0368	-0.4481	-0.3038

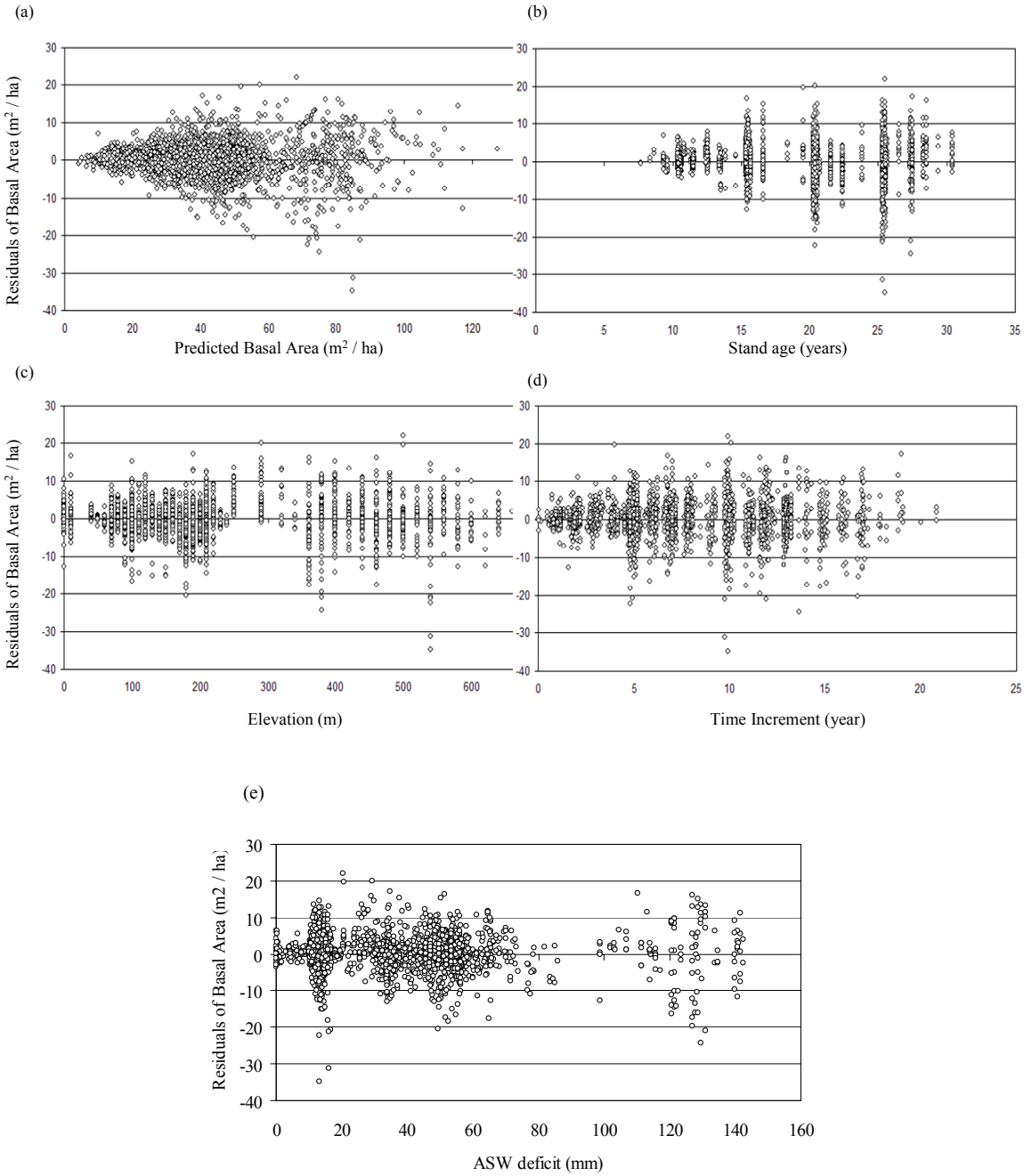


Figure 6.5: Fitting residual patterns of the final model for basal area: (a) residuals versus predicted, (b) residuals versus stand age, (c) residuals versus elevation, (d) residuals versus time increment (e) residuals versus ASW deficit .

Mean Top Height

The four models with the best initial fit for mean top height were Polymorphic Gompertz II, Polymorphic Von Bertalanffy-Richards II, Polymorphic Schumacher II, and the

Polymorphic Hossfeld as shown in Table 4.32. The effects of water balance deficit and elevation was added into these models by altering the asymptotic parameter for each equation. These equations were initially fitted with the form of elevation that was used for the final model of mean top height in CanSPBL(1.2) and the effects of water deficit added in as a linear and curvilinear term. The results of initial model fits in terms of MSE are shown in Table 6.6.

The model with the lowest MSE after the addition of the effect of average water deficit and elevation was altered further to see if different correction forms for these effects would improve model fits. These included a linear correction for elevation for PSP's above 250 m, as was used by Zhao (1999), elevation, the square of elevation, a probit correction for the effect of elevation, ASW deficit, ASW deficit squared, and the interaction of elevation and ASW deficit.

Table 6.6: Mean top height model fitting results, including the effect of average water deficit and elevation in the asymptotic parameter.

Equation with elevation and average 3PG water deficit in asymptote	Mean Top Height MSE
Polymorphic Gompertz II	1.01
Polymorphic Von Bertalanffy-Richards II	1.02
Polymorphic Schumachar II	0.97
Polymorphic Hossfeld	1.80

The best model fit was found by adjusting the asymptotic parameter of the Polymorphic Schumacher II equation using elevation squared, a dummy variable for elevations above 450m, mean ASW deficit, and mean ASW deficit squared. The final model for mean top height had a MSE of 0.97 and is written as:

$$MTH_2 = e^{\ln(MTH_1)\left(\frac{T_1+\gamma}{T_2+\gamma}\right)^{\beta} + \frac{(\alpha_0+\alpha_1\cdot elev^2+\alpha_2\cdot(elev-450)\cdot X+\alpha_3\cdot ASW+\alpha_4\cdot ASW^2)}{10000}\cdot\left(1-\left(\frac{T_1+\gamma}{T_2+\gamma}\right)^{\beta}\right)}$$

(6.6)

Where MTH_2 is the future mean top height (m), MTH_1 is the initial mean top height (m), T_1 is the initial stand age (years), T_2 is the final stand age (years), *elev* is the stand elevation (m), X is a binary indicator variable, where $X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450, ASW is the average water deficit over the simulation period calculated with the 3PG water balance model (mm), and β , γ, α_0, α_1, α_2, α_3, and α_4 are parameters whose values are listed in table 6.7.

The parameter estimates of the final model of mean top height are listed in table 6.7. The model showed no signs of bias for mean top height when plotted against prediction, age, elevation, and time increment (Figure 6.6).

Table 6.7: Parameters for mean top height model (Equation 6.6), standard errors, and approximate 95% confidence limits calculated with 3659 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.6832	0.0425	0.5998	0.7666
γ	4.1388	0.4782	3.2012	5.0764
α_0	44652.2	642.1	43393.3	45911.2
α_1	0.0238	0.00103	0.0218	0.0258
α_2	-32.8019	2.5632	-37.8276	-27.7763
α_3	40.0566	4.1723	31.8761	48.2371
α_4	-0.4519	0.039	-0.5285	-0.3754

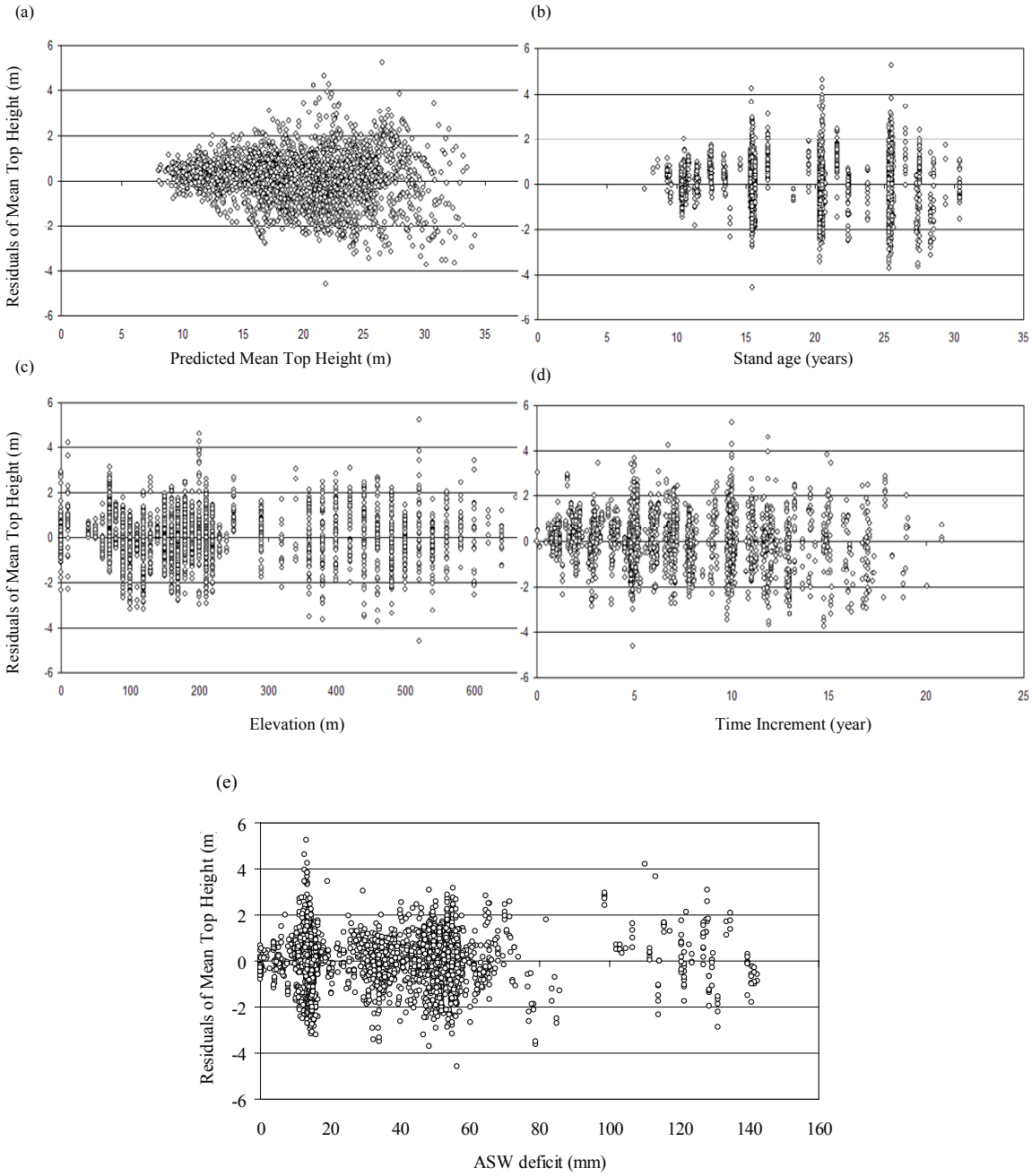


Figure 6.6: Fitting residuals patterns of final model for mean top height: (a) residuals versus predicted, (b) residuals versus stand age, (c) residuals versus elevation, and (d) residuals versus time increment, and (e) residuals versus ASW deficit.

Maximum diameter and standard deviation of diameter

Models for maximum diameter and standard deviation of diameter were fitted by selecting the 4 model forms that had the best initial fit for basal area. These forms were also fitted to the data, incorporating the effects of average water deficit and altitude

within the asymptotic parameter where appropriate. Results for the initial fitting of these models for maximum diameter and standard deviation of diameter are shown in Table 6.8.

Table 6.8: Mean square error for initial models of maximum diameter and standard deviation of diameter including the effect of average water deficit and elevation.

Equation with elevation and average 3PG water deficit in asymptote	Maximum Diameter	Standard Deviation of Diameter
Polymorphic Weibull I	4.89	0.39
Polymorphic Gompertz II	5.09	0.38
Polymorphic Weibull II	Failed to converge	Failed to converge
Polymorphic Schumacher II	5.06	0.43

The best model fit for maximum diameter was found by adjusting the asymptotic parameter of the Polymorphic Weibull I equation using elevation squared, a dummy variable for elevations above 450m, the inverse of initial stocking, and ASW deficit squared (Eq. 6.7). The best model fit for the standard deviation of diameter was similarly found by adjusting the asymptotic parameter for the Polymorphic Gompertz II equation using elevation squared, a dummy variable for elevations above 450m, ASW deficit, and ASW deficit squared (Eq. 6.8). The final models for maximum diameter and standard deviation of diameter had MSE values of 4.59 and 0.38 respectively and are written as:

$$DMAX_2 = DMAX_1 \cdot e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)} + \left(\alpha_0 + \alpha_1 \cdot elev^2 + \alpha_2 \cdot \left(elev - 450\right)\cdot X + \alpha_3 \cdot \frac{1}{N_1} + \alpha_4 \cdot ASW^2\right)\cdot\left(1 - e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)}\right)$$

(6.7)

$$DSTD_2 = e^{\left[\ln\left(DSTD_1\right)e^{\left(-\beta\cdot\left(T_2-T_1\right)+\gamma\left(T_2^2-T_1^2\right)\right)}\right]}\cdot e^{\left[\frac{\left(\alpha_0+\alpha_1*elev^2+\alpha_2\cdot\left(elev-450\right)\cdot X+\alpha_3\cdot ASW+\alpha_4\cdot ASW^2\right)}{10000}\left(1-e^{\left(-\beta\cdot\left(T_2-T_1\right)+\gamma\left(T_2^2-T_1^2\right)\right)}\right)\right]}$$

(6.8)

Where $DMAX_2$ is the future maximum diameter (m), $DMAX_1$ is the initial maximum diameter (m), $DSTD_2$ is the future standard deviation of diameter (m), $DSTD_1$ is the initial standard deviation of diameter (m),T_1 is the initial stand age (years), T_2 is the final stand age (years), *elev* is the stand elevation (m), X is a binary indicator variable, $X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450, ASW is the average water deficit over the simulation period calculated with the 3PG water balance model (mm), N1 is the initial stocking (stems / ha), and β , γ, α_0, α_1, α_2, α_3, α_4, and α_5 are parameters whose values are listed in table 6.9 (for maximum diameter eq. 6.6) and table 6.10 (for standard deviation of diameter eq. 6.7).

Tables 6.9, and 6.10 show parameter estimates. Figures 6.7, and 6.8 show residual patterns with little apparent bias against main modelling variables. For projection lengths of up to 20 years 95% of residuals for maximum diameter projections were within ± 4.3 cm, while 95% of the residuals for standard deviation of diameter were within ± 1.2 cm.

Table 6.9: Parameter values for maximum diameter model (Equation 6.6). Also shown are standard errors, and approximate 95% confidence limits for each parameter. Statistical values are calculated with 3659 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.3657	0.0421	0.2832	0.4483
γ	0.3973	0.0333	0.332	0.4626
α_0	64.8266	3.1145	58.7203	70.933
α_1	0.000159	0.00001	0.000139	0.000179
α_2	-0.1755	0.0176	-0.21	-0.141
α_3	8772.2	732.1	7336.8	10207.6
α_4	0.000463	0.000091	0.000284	0.000642

Table 6.10: Parameter values for standard deviation of diameter model (Equation 6.7). Also shown are standard errors, and approximate 95% confidence limits for each parameter. Statistical values are calculated with 3659 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.0715	0.00182	0.068	0.0751
γ	0.000443	0.000061	0.000323	0.000562
α_0	24777.1	358.5	24074.2	25480
α_1	0.0213	0.00109	0.0192	0.0235
α_2	-39.633	3.5957	-46.6829	-32.5831
α_3	-20.4624	6.0352	-32.2953	-8.6296
α_4	0.0935	0.0463	0.00276	0.1843

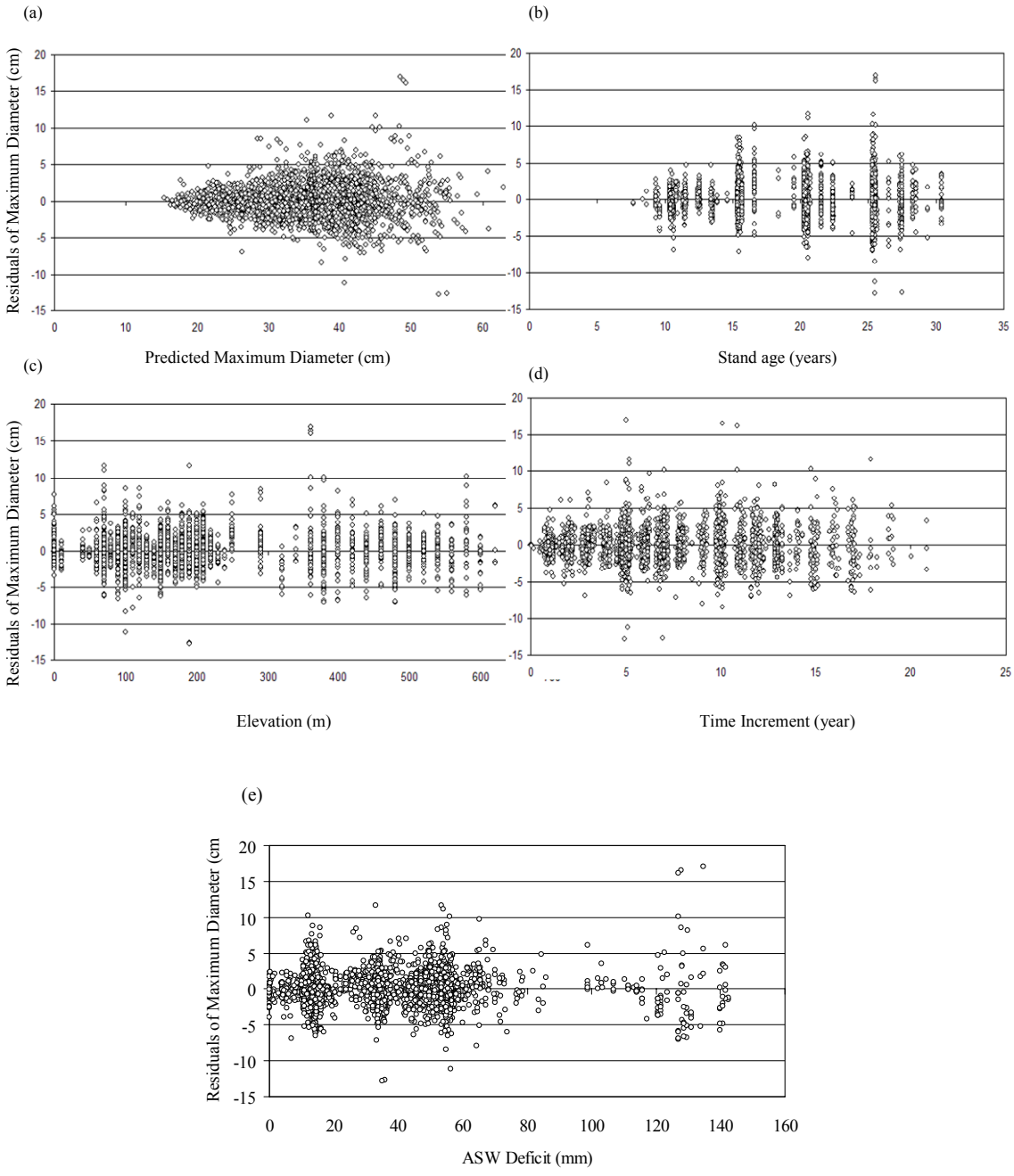


Figure 6.7: Residual patterns of final model for maximum diameter: (a) residuals versus predicted, (b) residuals versus stand age, (c) residuals versus elevation, (d) residuals versus time increment, and (e) residuals versus ASW Deficit.

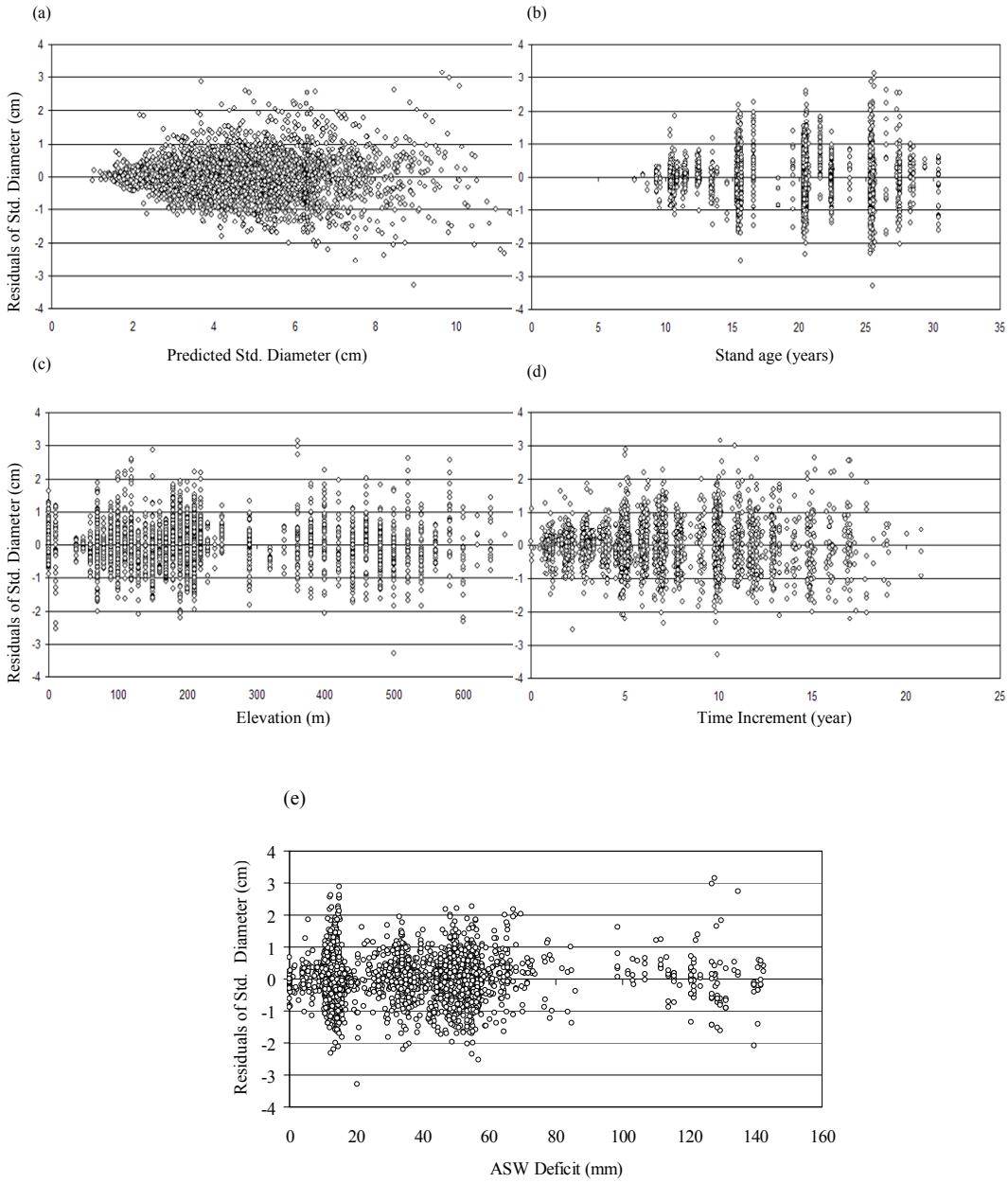


Figure 6.8: Residuals patterns of final model for standard deviation of diameter: (a) residuals versus predicted, (b) residuals versus stand age, (c) residuals versus elevation, (d) residuals versus time increment, and (e) residuals versus ASW deficit.

Mortality

A set of six equations which are variants of those listed by Clutter *et al.* (1983), those used by Zhao (1999), and those used to fit CanSPBL(1.2) were used as the basic equation forms for projecting stocking (Table 4.10). These equations were initially fitted using a modelling data set filtered in the same way as that used to fit CanSPBL(1.2) Equations were fitted incorporating the effects of ASW deficit and ASW deficit squared and the results of the initial fitting are listed in table 6.11. As equation 3 had the smallest MSE it was chosen as the basic form of difference equation for further analysis. This model was altered further to see if different correction forms for these effects would improve model fits. These included a linear correction for elevation for PSP's above 250 m, as was used by Zhao (1999), elevation, the square of elevation, a probit correction for the effect of elevation, ASW deficit, ASW deficit squared, and the interaction of elevation and ASW deficit.

Table 6.11: Difference equation forms for projection of stocking. Initial fitted models including ASW deficit and ASW deficit squared in asymptotic parameter. Original equation forms are listed in Table 4.10 of chapter 4.

Difference Equation Forms	Parameter Estimates		Model MSE
1: $N_2 = \left(N_1^{\,c} + \frac{a_0 + a_1 \cdot ASW + a_2 \cdot ASW^2}{10^6} \cdot \left(T_2^{\,b} - T_1^{\,b} \right) \right)^{(1/c)}$	a_0	6.2223	3285
	a_1	-0.0488	
	a_2	0.000802	
	b	1.8863	
	c	-0.5192	
2: $N_2 = \left(\frac{1}{\sqrt{N_1}} + \frac{a_0 + a_1 \cdot ASW + a_2 \cdot ASW^2}{10^8} \cdot \left(T_2^{\,b} - T_1^{\,b} \right) \right)^{-2}$	a_0	652.7	3283
	a_1	-5.6079	
	a_2	0.0844	
	b	1.8991	
3. $N_2 = \left(\frac{1}{\sqrt{N_1}} + \left(a_0 + a_1 \cdot ASW + a_2 \cdot ASW^2 \right) \cdot \left(\left(\frac{T_2}{100} \right)^2 - \left(\frac{T_1}{100} \right)^2 \right) \right)^{-2}$	a_0	0.0462	3282

	a_1	-0.00035	
	a_2	$5.828*10^{-6}$	
4. $N_2 = N_1 \cdot e^{\left(a_0+a_1\cdot ASW+a_2\cdot ASW^2\right)\left(T_2-T_1\right)}$	a_0	-868.5	3488
	a_1	10.4256	
	a_2	-0.2191	
5. $N_2 = N_1 \cdot \left(\frac{T_2}{T_1}\right)^{\left(a_0+a_1\cdot ASW+a_2\cdot ASW^2\right)} \cdot e^{b\cdot\left(T_2-T_1\right)}$	a_0	0.1773	3328
	a_1	0.00283	
	a_2	-0.00004	
	b	-0.0217	
6. $N_2 = N_1 \cdot e^{\left(a_0+a_1\cdot ASW+a_2\cdot ASW^2\right)\left(T_2^{\,b}-T_1^{\,b}\right)}$	a_0	-18.2851	3391
	a_1	0.1299	
	a_2	-0.00310	
	b	2.0982	

The final model selected (Eq. 6.9) was fitted using Equation 3 (Table 4.10) and produced a MSE of 3282. Parameters for the final mortality model (Eq. 6.9) are listed in Table 6.12. Little bias was apparent in residual patterns shown in Figure 6.9. Ninety five percent (>2.5% and <97.5%) of residuals were within ± 116 stems / ha for projection lengths of up to 21 years.

$$N_2 = \left(\frac{1}{\sqrt{N_1}} + \left(\alpha_0 + \alpha_1 \cdot ASW + \alpha_2 \cdot ASW^2\right)\cdot\left(\left(\frac{T_2}{100}\right)^2 - \left(\frac{T_1}{100}\right)^2\right)\right)^{-2} \qquad (6.9)$$

Where N_2 is the future stocking (stems / ha), N_1 is the initial stocking (stems / ha), T_1 is the initial stand age (years), T_2 is the final stand age (years), ASW is the average water deficit over the simulation period calculated with the 3PG water balance model (mm), and α_0, α_1, and α_2 are parameters whose values are listed in table 6.12 (for stocking eq. 6.8).

Table 6.12: Stocking model parameters for equation 6.9, standard errors, and approximate 95% confidence limits calculated with 2226 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
α_0	0.0462	0.00313	0.0401	0.0524
α_1	-0.00035	0.000119	-0.00058	-0.00012
α_2	5.83E-06	8.13E-07	4.23E-06	7.42E-06

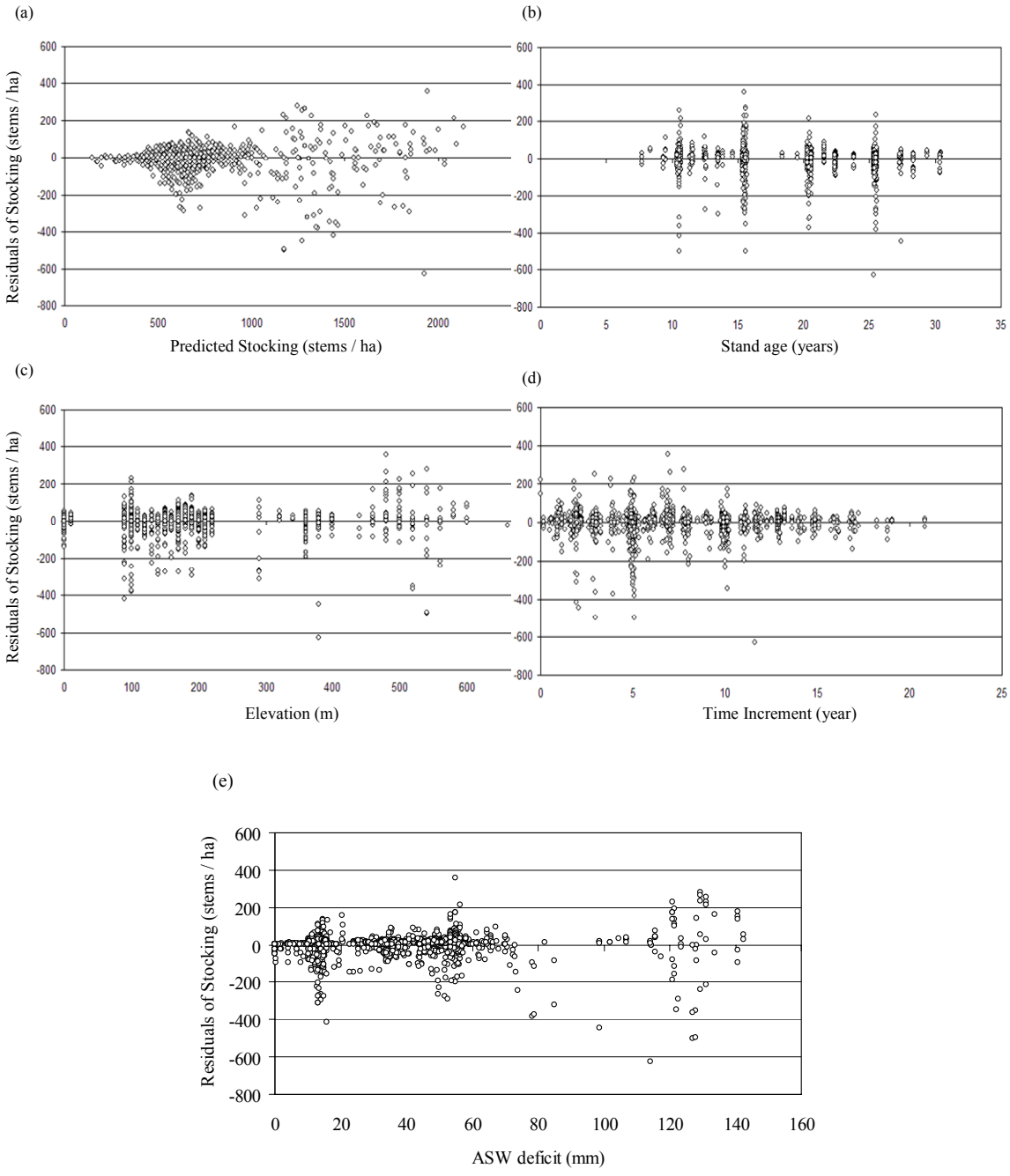


Figure 6.9: Residuals patterns of final model for stocking: (a) residuals versus predicted, (b) residuals versus stand age, (c) residuals versus elevation, (d) residuals versus time increment, and (e) residuals versus ASW deficit.

Testing Model Parameters

Final model parameters were tested against an autocorrelation-free dataset to see if they were significantly different from zero. An autocorrelation-free dataset was prepared from the original modelling data by randomly selecting one observation from each of the 746 plots. Parameters for the models of mean top height, basal area, maximum diameter, standard deviation of diameter, and stocking were all found to be significant at the 95% confidence level (Tables 6.13, 6.14, 6.15, 6.16, and 6.17).

Table 6.13: Basal area model parameters for equation 6.5, standard errors, and approximate 95% confidence limits tested with an autocorrelation-free dataset with 745 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.153	0.00419	0.1448	0.1612
γ	0.00241	0.00008	0.00226	0.00257
α_0	43663.9	451.9	42776.7	44551
α_1	0.0355	0.00147	0.0326	0.0383
α_2	-36.2397	4.3294	-44.7393	-27.7401
α_3	56.5576	10.2484	36.4378	76.6774
α_4	-0.5827	0.0901	-0.7596	-0.4058

Table 6.14: Mean top height model parameters for equation 6.6, standard errors, and approximate 95% confidence limits tested with an autocorrelation-free dataset with 745 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	8.1675	1.6642	4.9002	11.4347
γ	1.0427	0.1488	0.7506	1.3349
α_0	41481	965	39586.6	43375.4
α_1	0.0214	0.00177	0.018	0.0249
α_2	-39.6227	4.5143	-48.4852	-30.7601
α_3	22.6608	8.0674	6.8228	38.4989
α_4	-0.2664	0.0723	-0.4083	-0.1246

Table 6.15: Maximum diameter model parameters for equation 6.7, standard errors, and approximate 95% confidence limits tested with an autocorrelation-free dataset with 745 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.7474	0.2499	0.2568	1.2381
γ	0.2681	0.0761	0.1187	0.4175
α_0	77.1989	9.9047	57.7539	96.644
α_1	0.000192	0.000031	0.00013	0.000253
α_2	-0.2677	0.0515	-0.3688	-0.1665
α_3	0.00127	0.000349	0.000587	0.00196

Table 6.16: Standard deviation of diameter model parameters for equation 6.8, standard errors, and approximate 95% confidence limits tested with an autocorrelation-free dataset with 745 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.0649	0.00372	0.0576	0.0722
γ	0.000382	0.000122	0.000143	0.000622
α_0	26173.8	988.9	24232.3	28115.2
α_1	0.0232	0.00247	0.0183	0.028
α_2	-59.269	8.4106	-75.7807	-42.7573
α_3	-40.3089	15.4589	-70.6581	-9.9598
α_4	0.2613	0.1126	0.0402	0.4823

Table 6.17: Mortality model parameters for equation 6.9, standard errors, and approximate 95% confidence limits tested with an autocorrelation-free dataset with 745 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
α_0	0.047	0.00567	0.0359	0.0582
α_1	-0.00048	0.000214	-0.0009	-0.00006
α_2	9.87E-06	1.58E-06	6.77E-06	1.30E-05

6.6: Model Validation and Comparison with CanSPBL(1.2)

Validation of CanSPBL(water) and CanSPBL(1.2) with a subset of the validation data set presented previously in chapter 4, produced model fitting statistics and residual distribution statistics listed in Tables (6.18) and Table (6.19) respectively. CanSPBL(water) showed improvements for most model components of 1.07 to 3.77% after being updated with an index of root zone water balance. However, CanSPBL(water) showed a -0.78 % decrease in precision for the component of maximum diameter. CanSPBL(water) also showed less bias in prediction of stocking, and basal area (Table 6.18). Residual distributions for CanSPBL(water) showed a worse fit overall for skewness and kurtosis, while the hypothesis of normality of residuals was not rejected for both models with the Kolmogorov-Smirnov test (Table 6.19).

Table 6.18: Comparison of model statistics between CanSPBL(1.2), and CanSPBL(water). Statistics shown include Mean Square Error (MSE), Average Model Bias (AMB), Model Efficiency Factor (EF), and percent difference of MSE (% Difference of MSE) n = 967.

Model		Model Fitting Statistics			
		MSE	AMB	EF	MSE % Difference
CanSPBL(1.2)					
	Stocking	2154.92	6.51	0.95	2.84
	Basal Area	17.93	0.55	0.96	3.77
	Mean Top Height	1.10	0.08	0.96	2.95
	Maximum Diameter	5.60	0.27	0.93	-0.78
	Standard Deviation of Diameter	0.53	0.04	0.86	1.07
CanSPBL2(water)					
	Stocking	2094.63	1.96	0.96	-2.84
	Basal Area	17.27	0.45	0.96	-3.77
	Mean Top Height	1.07	0.05	0.96	-2.95
	Maximum Diameter	5.64	0.27	0.93	0.78
	Standard Deviation of Diameter	0.52	0.04	0.86	-1.07

Table 6.19: Model normality distribution statistics for CanSPBL(1.2), and CanSPBL(water) in terms of skewness, kurtosis, and p-value from the Kolmogorov-Smirnov test for normality n = 967.

Model		Residual Distribution Statistics		
		Skewness	Kurtosis	P-value
CanSPBL(1.2)				
	Stocking	-1.01	11.05	0.01
	Basal Area	0.42	3.74	0.01
	Mean Top Height	0.13	0.96	0.01
	Maximum Diameter	1.15	4.73	0.01
	Standard Deviation of Diameter	1.22	7.33	0.01
CanSPBL2(water)				
	Stocking	-2.80	14.02	0.01
	Basal Area	0.22	4.05	0.01
	Mean Top Height	-0.02	1.10	0.01
	Maximum Diameter	1.19	4.84	0.01
	Standard Deviation of Diameter	1.20	7.35	0.01

A separate validation was also completed with the same models using long term average climate as inputs into the water balance model. This was done to test how model accuracy might be affected if managers were to estimate future climate using average values. This separate validation of CanSPBL(1.2), and CanSPBL(water_using average climate) produced model fitting statistics and residual distribution statistics listed in Tables (6.20) and Table (6.21) respectively. CanSPBL(water_using average climate) showed improvements for most model components of 0.47 to 7.81% after being updated with an index of root zone water balance that uses average climate inputs. However, CanSPBL(water_using average climate) showed a -0.13 decrease in precision for the component of maximum diameter. These results indicate that model accuracy for the prediction of stocking, basal area, and maximum diameter would actually increase precision using average climate inputs as opposed to those that describe yearly variations. CanSPBL(water_using average climate) showed less bias in prediction of stocking, basal area, and maximum diameter than did CanSPBL(water) (Table 6.18 and Table 6.20). Residual distributions for CanSPBL(water_using average climate) showed a worse fit overall for skewness and kurtosis in comparison to CanSPBL(1.2), with the exceptions of skewness for basal area and mean top height. The hypothesis of normality of residuals was not rejected for CanSPBL(water_using average climate) with the Kolmogorov-Smirnov test (Table 6.21). Although the distributions of residuals for CanSPBL(water_using average climate) are better on average than CanSPBL(water).

Table 6.20: Comparison of model statistics between CanSPBL(1.2), and CanSPBL(water) (using average climate inputs). Statistics shown include Mean Square Error (MSE), Average Model Bias (AMB), Model Efficiency Factor (EF), and percent difference of MSE (% Difference of MSE) n = 967.

Model		Model Fitting Statistics			
		MSE	AMB	EF	MSE % Difference
CanSPBL(1.2)					
	Stocking	2154.92	6.51	0.95	7.81
	Basal Area	17.93	0.55	0.96	3.82
	Mean Top Height	1.10	0.08	0.96	2.57
	Maximum Diameter	5.60	0.27	0.93	-0.13
	Standard Deviation of Diameter	0.53	0.04	0.86	0.47
CanSPBL2(water) Using average climate inputs					
	Stocking	1992.87	3.11	0.96	-7.81
	Basal Area	17.26	0.51	0.96	-3.82
	Mean Top Height	1.07	0.08	0.96	-2.57
	Maximum Diameter	5.60	0.24	0.93	0.13
	Standard Deviation of Diameter	0.52	0.05	0.86	-0.47

Table 6.21: Model normality distribution statistics for CanSPBL(1.2), and CanSPBL(water) (using average climate inputs) in terms of skewness, kurtosis, and p-value from the Kolmogorov-Smirnov test for normality n = 967.

Model		Residual Distribution Statistics		
		Skewness	Kurtosis	P-value
CanSPBL(1.2)				
	Stocking	-1.01	11.05	0.01
	Basal Area	0.42	3.74	0.01
	Mean Top Height	0.13	0.96	0.01
	Maximum Diameter	1.15	4.73	0.01
	Standard Deviation of Diameter	1.22	7.33	0.01
CanSPBL2(water) Using average climate inputs				
	Stocking	-2.57	12.53	0.01
	Basal Area	0.27	4.12	0.01
	Mean Top Height	0.02	1.04	0.01
	Maximum Diameter	1.16	4.73	0.01
	Standard Deviation of Diameter	1.25	7.62	0.01

6.7: Discussion

Incorporating ASW deficit into equations that predict basal area, mean top height, standard deviation of diameter, and stocking in the growth and yield model CanSPBL(1.2) showed a small, improvement in model accuracy (Table 6.18). There was somewhat of an increased precision for longer intervals shown in residual plots (Figures 6.5, 6.6, 6.7, and particularly Figure 6.9). This effect may have been due to the levelling of water balance models precision increasing with time of simulation. If initial water balance estimates are different from measurements, they become more precise over time for months following initialization, and this may affect growth and yield model precision for shorter measurement intervals. The results of this study show a relatively small improvement in model accuracy and growth pattern including ASW deficit into the models. This result was surprising in the context of varying climate over the study area. With yield modelling the results could have been markedly more significant, as the growing condition of the site is not implicit in initial stem size (as is used in the

difference approach). This result highlights the power of the difference modelling approach within and between regions with varying climatic and site conditions.

Validation with average climatic inputs surprisingly increased precision of the models that predict stocking, basal area, and maximum diameter. Predicted stocking showed the largest reduction of residual mean square error, at 7.81% using average climate. This result is positive for forest managers who may only have average climate available to run such models. As regions vary in climate, taking this into account can result in small gains in accuracy.

The updated version of the model included both ASW deficit and a simplified form of elevation. This result indicates that the effect of elevation in the Canterbury region describes more than water relations on the site and may be a proxy for other site factors such as nutrient availability, or possibly average temperature.

The data required to run the water balance model are monthly values of maximum temperature, minimum temperature, solar radiation, vapour pressure deficit, leaf area index, rainfall, soil type, initial estimates of available soil water, maximum available soil water, and minimum available soil water. Obtaining these inputs for a given forest can be time consuming and expensive and may not warrant the 1 – 8% improvement in accuracy that these models offer. There is also the issue of even lower precision for the models of mean top height and standard deviation of diameter using average values of climatic inputs to describe future conditions where monthly and yearly variation may not be available. The increase in model accuracy for the models of mean top height and standard deviation of diameter using inputs that including monthly and yearly variation are 2.95 and 1.07% respectively. There is a decrease in precision for the models of mean top height and standard deviation of diameter to 2.57 and 0.47% respectively using average monthly climatic inputs. These results suggest model accuracy may be maximised by using average monthly inputs for some models and including monthly and yearly variation for the inputs of the models of mean top height and standard deviation of diameter. Model precision could be increased further by considering the effect of drought only in months where growth is likely to occur. Drought in the winter months may not have the same effect on growth as drought that occurred during the growing season. Future research in this area could focus on trying to add a weight to ASW deficits

according to the month in which they occur.

6.8: Conclusions

Model components of mean top height, basal area per hectare, stems per hectare, and diameter distribution were developed by using non linear least squares regression techniques to select appropriate equation forms. Explanatory variables to improve model prediction were tested and incorporated into models where appropriate. The effect of elevation, and ASW deficit was added to models of mean top height, basal area, and diameter distribution, only the effect of ASW deficit was added to the model of mortality. A polymorphic Gompertz equation displayed the best fit for basal area, while a polymorphic Schumacher equation displayed the best fit for mean top height. The diameter distribution model of maximum diameter displayed the best fit with a polymorphic Weibull equation, while the standard deviation of diameter model fit best with a polymorphic Gompertz equation.

A mortality severity index based on the -3/2 power law was used as a basis to filter the modelling data set for mortality and a final model was built with 80% of the original mortality records under the recommendation of managers at SPBL. Residual mean square error for the model of stocking was decreased by 2.84% by incorporating ASW deficit calculated by the equation (6.10). Residual mean square error was reduced further by 7.8% by using average climatic inputs. Ninety five percent (>2.5% and <97.5%) of residuals were within ± 116 stems / ha for projection lengths of up to 21 years using monthly climatic inputs that show yearly variation, and within ± 92 stems / ha using average climatic inputs. Future stocking is dependent on the current stocking (stems / ha), initial stand age (years), final stand age (years), elevation squared (m^2), ASW deficit (mm), and ASW deficit squared (mm^2).

$$N_2 = \left(\frac{1}{\sqrt{N_1}} + \left(\alpha_0 + \alpha_1 \cdot ASW + \alpha_2 \cdot ASW^2\right) \cdot \left(\left(\frac{T_2}{100}\right)^2 - \left(\frac{T_1}{100}\right)^2 \right) \right)^{-2} \qquad (6.10)$$

All model parameters were tested against an auto-correlation free dataset for significance. Tests showed that all parameters were significant at a 95% confidence level. The entire updated version of the model CanSPBL(water) was validated against an independent data

set and compared to a the original CanSPBL(1.2). CanSPBL(water) showed improvements in MSE of 1 to 3% after the effects of ASW deficit were incorporated into the model, however the model of maximum diameter showed a worse fit by 0.78%.

Residual distributions for CanSPBL(water) showed a worse fit overall for skewness and kurtosis. The hypothesis of normality was not rejected for both models using the Kolmogorov-Smirnov test.

Residual mean square error for the model of basal area was reduced by 3.77% by incorporating ASW deficit calculated by the equation (6.11). Ninety five percent (>2.5% and <97.5%) of residuals were within ± 8.7 m^2 / ha for projection lengths of up to 21 years using monthly climatic inputs that show yearly variation, and within ± 8.8 m^2 / ha using average climatic inputs. Future basal area is dependent on the current stand basal area (m^2 / ha), initial stand age (years), final stand age (years), elevation squared (m^2), a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450), ASW deficit (mm), and ASW deficit squared (mm^2).

$$G_2 = e^{\left(\ln(G_1) e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)} + \frac{\left(\alpha_0+\alpha_1\cdot elev^2+\alpha_2\cdot((elev-450)\cdot X)+\alpha_3\cdot ASW+\alpha_4\cdot ASW^2\right)}{10000}\cdot\left(1-e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right)\right)}$$

(6.11)

Residual mean square error for the model of mean top height was reduced by 2.95% by incorporating ASW deficit calculated by the equation (6.12). Ninety five percent (>2.5% and <97.5%) of residuals were within ± 2.1 m for projection lengths of up to 21 years using monthly climatic inputs that show yearly variation, and within ± 2.2 m using average climatic inputs. Future mean top height is dependent on the current mean top height (m), initial stand age (years), final stand age (years), elevation (m), elevation squared (m^2), a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450), ASW deficit (mm), and ASW deficit squared (mm^2).

$$MTH_2 = e^{\ln(MTH_1)\left(\frac{T_1+\gamma}{T_2+\gamma}\right)^{\beta} + \frac{\left(\alpha_0+\alpha_1\cdot elev^2+\alpha_2\cdot(elev-450)\cdot X+\alpha_3\cdot ASW+\alpha_4\cdot ASW^2\right)}{10000}\cdot\left(1-\left(\frac{T_1+\gamma}{T_2+\gamma}\right)^{\beta}\right)}$$

(6.12)

Residual mean square error for the model of maximum diameter was increased by 0.78% by incorporating ASW deficit calculated by the equation (6.13). Ninety five percent (>2.5% and <97.5%) of residuals were within ± 5 cm for projection lengths of up to 21 years using monthly climatic inputs that show yearly variation, and within ± 5 cm using average climatic inputs. Future maximum diameter is dependent on the current maximum diameter (cm), initial stand age (years), final stand age (years), elevation squared (m^2), a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450), the inverse of initial stocking (ha / stems), and ASW deficit squared (mm^2).

$$DMAX_2 = DMAX_1 \cdot e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)} + \left(\alpha_0 + \alpha_1 \cdot elev^2 + \alpha_2 \cdot (elev - 450)\cdot X + \alpha_3 \cdot \frac{1}{N_1} + \alpha_4 \cdot ASW^2\right)\cdot\left(1 - e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)}\right) \quad (6.13)$$

Residual mean square error for the model of standard deviation of diameter was decreased by 1.07% by incorporating ASW deficit calculated by the equation (6.14). Ninety five percent (>2.5% and <97.5%) of residuals were within ± 1.5 cm for projection lengths of up to 21 years using monthly climatic inputs that show yearly variation, and within ± 1.5 cm using average climatic inputs. Future standard deviation of diameter is dependent on the current standard deviation of diameter (cm), initial stand age (years), final stand age (years), elevation squared (m^2), a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450), ASW deficit (mm), and ASW deficit squared (mm^2).

$$DSTD_2 = e^{\left[\ln(DSTD_1)e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right]}\cdot e^{\left[\frac{\left(\alpha_0+\alpha_1*elev^2+\alpha_2\cdot(elev-450)\cdot X+\alpha_3\cdot ASW+\alpha_4\cdot ASW^2\right)}{10000}\cdot\left(1-e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right)\right]} \quad (6.14)$$

CHAPTER 7

COMPARISON OF MODELLING APPROACHES

7.1 Introduction

Over the past decade there has been much debate within the scientific community on the relative merits of modelling forest growth with empirical-statistical versus process-based approaches. This chapter aims to make an objective comparison and validation of a range of model types, with the main criterion for comparison being each model's ability to match actual historical measurements of forest growth in an independent data set. The models compared were CANTY (Goulding, 1995), CanSPBL(1.2) (Pinjuv, 2005), CanSPBL-water (Pinjuv, 2005), and 3-PG (Landsberg and Waring, 1997). CANTY (Goulding, 1995), is an existing empirical stand level growth and yield model for radiata pine plantations in Canterbury. CanSPBL(1.2) (Pinjuv, 2005), is an updated version of CanSPBL (Zhao, 1999), a stand-level growth and yield model for radiata pine plantations growing in Canterbury. CanSPBL-water (Pinjuv, 2005) is an updated version of CanSPBL (Zhao, 1999) with an index for root zone water balance over the growth interval added to the model. Finally, 3-PG (Landsberg and Waring, 1997), is a simplified process-based growth model that can be adapted to a range of forest species by parameterisation of model coefficients.

7.2 Methods

The models were examined quantitatively, assessing model behaviour with a validation data set. The procedure involved graphical displays and statistical tests. Potential correlation was detected with inspection of graphical plots of residual versus predictions and explanatory variables. Model residual errors or the difference between predictions made with the various models and observations of growth with the current data will display certain trends along with initial conditions, projection length, or predicted values when the model is biased. 3PG was the only model evaluated that required site specific

calibration before it could be validated, though it can be argued that fitting sigmoid equations to local PSP datasets may also represent local parameterisation. 3PG was calibrated using an independent data set and varying select input parameters of the model to give good fits between measured and modelled variables.

7.21 Description of Models

Overview of CANTY

The statistical growth and yield model CANTY is a stand level model including the components of mean top height, basal area/ha, stems/ha and volume/ha. The model was built by the New Zealand Forest Research Institute (1991) and was intended for modelling growth and yield of radiata pine growing in the Canterbury region (Goulding, 1995). CANTY is a state-space model, were a state is defined by several state vectors. The state of a stand in such models is usually expressed by mean top height, basal area, stems per hectare, and sometimes crown closure. Future states can be predicted by future management options and current states that are summaries of past growth. The behaviour of the system is described by a transition function and an output function. The output function for volume is predicted with state variables. The state space approach is critically dependent on site index, which is estimated with a height function, which in turn can be highly dependent on the measurement intervals used in model fitting.

Some theoretical aspects of the model are described by the following;

Transition function: $X(t) = F[X(t_0), U, T-T_0]$ or $dX/dt = f(X)$ (7.1)

Output Function: $V(t) = G[X(t)]$, or $dV/dt = g(X)$ (7.2)

Where: X = (H,G,N), H = mean top height, G = basal area, N = stocking, V = volume, U = input (management options) T_0 = starting time of a period, T = the end of a period of time

When a multivariate generalisation of the Bertlanffy-Richards model is adopted, the new state vectors become:

$$Y = (y_1, y_2, y_3) = \left(H^{C_{11}}, H^{C_{21}} G^{C_{22}} N^{C_{23}}, H^{C_{31}} G^{C_{32}} N^{C_{33}}\right) \text{ or } Y = X^C. \quad (7.3)$$

The linear differential equation is:

$dY/dt = A\ Y + b = A\ X^C + b$ or

$dY/dt = A(X^C - a)$ when $a = A^{-1}\ b$ (7.4)

where A and C is a 3x3 matrix, a and b is a 1x3 matrix, $T = b_h t$ is scaled time, b_h is the coefficient in the site index equation.

A global difference equation

$$X(t_2) = \left\{a + P^{-1} e^{\Lambda b_h (t_2 - t_1)} P\left[X^c(t_1) - a\right]\right\}^{(1/c)} \quad (7.5)$$

is obtained from integration of the above differential equation. P and Λ are such that Λ is diagonal and $A = P^{-1}\Lambda P$.

A maximum likelihood estimator is used to estimate all equations simultaneously. The simultaneous estimation of all model components minimises overall errors but restricts the choice of functional form for individual state variables.

The model was prepared for testing by using a Basic code version provided by the Growth Modelling Research Cooperative. The code was adapted by Dr E.G. Mason so that it could work in batch mode.

Overview of 3-PG

The hybrid growth model 3-PG (Physiological Principles Predicting Growth) developed by Landsberg and Waring (1997), is a simplified process-based, stand level model of forest growth. This model expresses gross primary productivity (GPP) as the product of radiation-use efficiency (ε) and absorbed photosynthetically active radiation (APAR) (Equation 7.6). A set of modifiers reduce the efficiency of a unit of radiation as a result of soil water deficit (f_θ), vapour pressure deficit of the air (f_D), temperature (f_T), soil

nutrition (f_N), and stand age (f_A) (Landsberg and Waring, 1997). Where only the minimum of the f_N and the f_D modifiers is used. Net primary production (NPP) is calculated as a fixed amount of GPP (*c*) and the general equation can be expressed as:

$$NPP = GPP \cdot c = \left(\varepsilon \sum_{t-1}^{T} APAR_t \cdot f_\theta \cdot f_D \cdot f_T \cdot f_N \cdot f_A \right) \cdot c \qquad (7.6)$$

The 3-PG model consists of five simple process-based sub models: the assimilation of carbohydrates, the distribution of biomass between foliage, roots and stems, the determination of stem number, soil water balance, and the conversion of biomass values into variables of interest to forest managers.

The model is run on a monthly time step and the state of the stand is updated every month over the simulation period. The model requires input parameters that describe the growth characteristics of the species, site characteristics, and climatic inputs. The 3-PG model predicts the time-course of stand development, water use, and available soil water. Its primary output variables are net primary production, the standing biomass in foliage, stem, and roots, stem numbers, available soil water, and transpiration. The model also outputs leaf area index, mean stem diameter at breast height, main stem volume, and mean annual increment. The tested model (version 2.3) was provided by the Commonwealth Scientific and Industrial Research Organisation as Basic code within a Microsoft Excel spreadsheet.

Overview of CanSPBL(1.2)

The model CanSPBL(1.2) is a non linear least squares regression system of equations to predict forest growth. The modelling approach used was to create a stand level model. Predictions made by the model are mean top height (MTH), basal area per hectare, stems per hectare, volume per hectare and diameter distribution. Equations for each predicted variable were fitted using non-linear least-square regression procedures as described in chapter 4 (Pinjuv 2005). Diameter class distributions are described in the model using a reverse Weibull function. Stand tables can be produced with CanSPBL(1.2) using a recovery method of parameters to project future stand statistics, and finally the method of

moments is used to convert stand statistics of standard deviation, maximum value, and arithmetic mean diameter to Weibull distribution parameters.

Overview of CanSPBL(water)

The model CanSPBL(water) is a non linear least squares regression system of equations to predict forest growth that includes the effects of available soil water over the simulation period. Available soil water predictions within CanSPBL(water) are estimated with the sub-model for water balance adapted from the process-based growth model 3PG (Landsberg and Waring, 1997). The modelling approach used was to create a stand level model. Predictions made by the model are mean top height (MTH), basal area per hectare, stems per hectare, volume per hectare and diameter distribution. Equations for each predicted variable were fitted using non-linear least-square regression procedures as described in chapter 6 (Pinjuv 2005). Diameter class distributions are described in the model using a reverse Weibull function. Stand tables can be produced with CanSPBL(water) using a recovery method of parameters to project future stand statistics, and finally the method of moments is used to convert stand statistics of standard deviation, maximum value, and arithmetic mean diameter to Weibull distribution parameters. Both CanSPBL1.2 and CanSPBL(water) were implemented in Microsoft Excel as code written for the purpose of independent validation.

7.22 Description of the Validation Dataset

The validation dataset comprised 969 plot measurements taken within 195 plots. All validation data were independent of any information used for fitting or calibrating any of the included models in this study. A summary of plot variable estimates with a breakdown for plains and foothills plots is listed in Table 7.1, and Figure 7.1 displays growth patterns of the main variables. These revealed the following characteristics.

Table 7.1: Summary of data for model validation. Data shown represents 728 plot measurements taken from plots located in the plains, and 241 plot measurements from the hills.

Variable: **Plains Data**	Units	Mean	Standard Deviation	Minimum	Maximum
Age (Initial)	Years	11.5	3.8	7.5	25.5
Age (Final)	Years	18.0	5.2	7.7	30.4
Stocking (Initial)	Stems / ha	668.1	143.0	275.0	1600.0
Stocking (Final)	Stems / ha	645.0	137.4	275.0	1600.0
Mean Top Height (Initial)	m	12.4	3.7	7.2	25.9
Mean Top Height (Final)	m	19.1	4.9	8.7	29.3
Basal Area (Initial)	m^2 / ha	19.1	11.4	4.6	61.6
Basal Area (Final)	m^2 / ha	35.1	13.6	7.2	73.1
Dbh Std. D (Initial)	cm	2.8	1.2	0.8	8.2
Dbh Std. D (Final)	cm	4.4	1.5	0.9	8.7
Dbh Max. (Initial)	cm	23.4	6.3	13.0	48.2
Dbh Max. (Final)	cm	33.4	7.2	15.7	50.5
Elevation	m	135.8	60.1	0.0	250.0
Variable: **Hills Data**	Units	Mean	Standard Deviation	Minimum	Maximum
Age (Initial)	Years	12.9	4.6	7.6	25.5
Age (Final)	Years	19.6	5.7	9.6	29.4
Stocking (Initial)	Stems / ha	823.2	424.0	250.0	2100.0
Stocking (Final)	Stems / ha	754.7	362.7	250.0	1975.0
Mean Top Height (Initial)	m	13.0	5.2	6.5	30.8
Mean Top Height (Final)	m	21.1	6.8	8.4	33.9
Basal Area (Initial)	m^2 / ha	34.7	26.9	4.9	104.0
Basal Area (Final)	m^2 / ha	59.7	27.0	9.2	123.0
Dbh Std. D (Initial)	cm	4.1	1.8	1.9	12.2
Dbh Std. D (Final)	cm	6.3	2.4	2.2	13.1
Dbh Max. (Initial)	cm	29.6	9.6	16.5	63.6
Dbh Max. (Final)	cm	43.4	10.0	17.6	69.3
Elevation	m	455.6	85.6	290.0	660.0

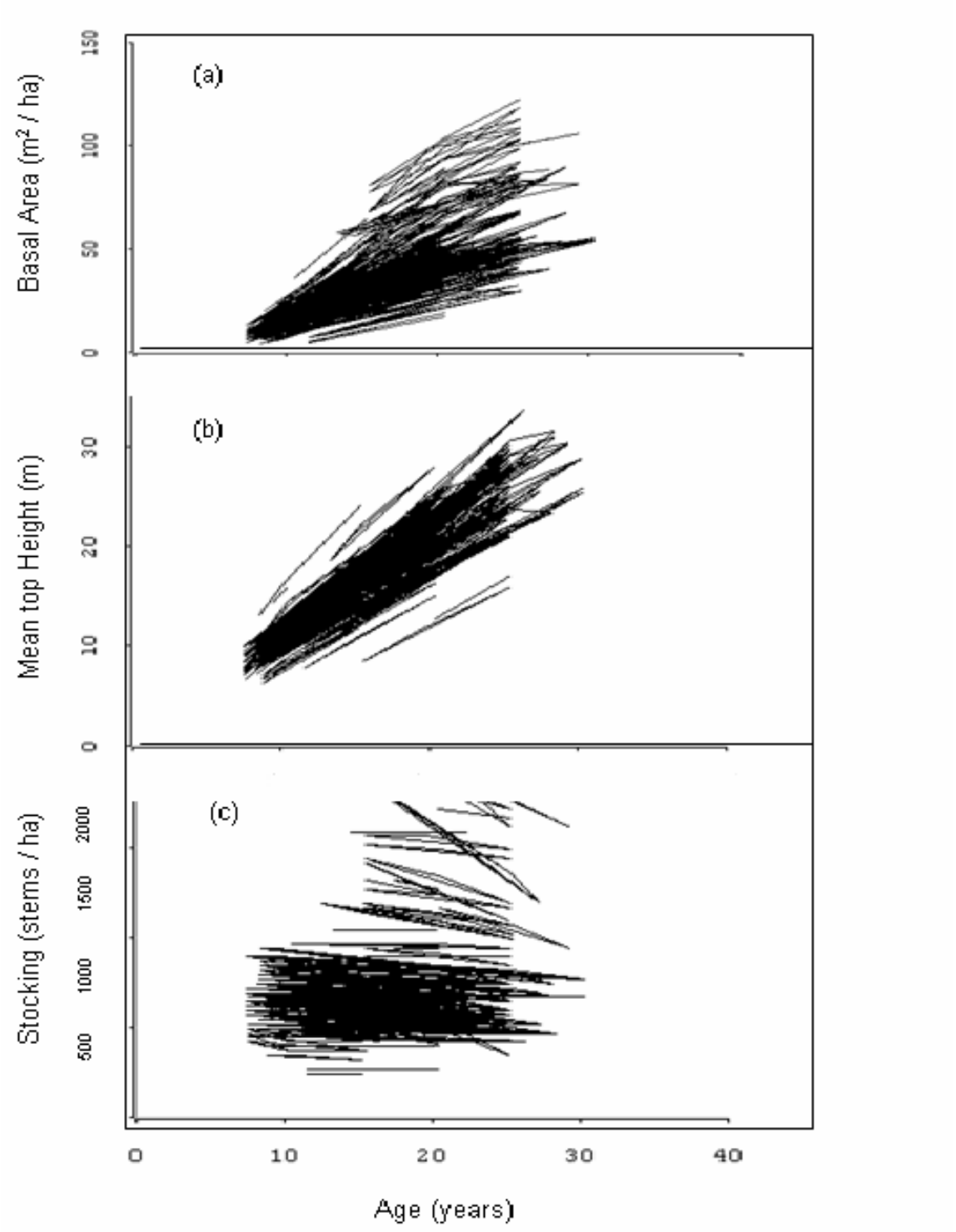


Figure 7.1: Growth patterns over time for (a) basal area, (b) mean top height, (c) stems per hectare using data for the whole model validation data set.

7.23 Validation Tests

Model validation procedures in this study involved assessing model behaviour with the validation data set. Models were examined quantitatively using graphical displays and statistical tests of Average Model Bias (AMB) and Model Efficiency (EF) (Loague and Green, 1991). Potential correlation was detected with inspection of graphical plots of residual versus predictions and explanatory variables for each of the main model components (mean top height, basal area, and stocking). Model residual errors or the difference between predictions made with final model equations and observations were examined for trends that indicated bias. Model performance has been described using AMB and EF (Loague and Green, 1991; Zhao, 1999). Average model bias (Equation 7.7) is an average of errors for all predictions. An AMB of 0 would indicate a model with no bias. Model efficiency (Equation 7.8) is also a measure of model performance. A high value of EF (maximum value of 1) indicates a model of perfect fit, while EF value of 0 indicates a model of poor fit where the average value would model the relationship as well. A negative EF value indicates an even poorer fit than the average value.

$$AMB = \frac{1}{n}\Sigma\left(Y_i - \hat{Y}_i\right) \tag{7.7}$$

$$EF = 1 - \frac{\Sigma\left(Y_i - \hat{Y}_i\right)^2}{\Sigma\left(Y_i - \overline{Y}_i\right)^2} \tag{7.8}$$

Where: Y_i observed, $\hat{Y}_i$ is the modelled, and $\overline{Y_i}$ is the average observed value.

Dependent variables tested for bias and fit were chosen from those outputs which were common among models, and those that were viewed as useful to forest managers. These included mean top height (MTH), basal area (G), and the number of stems per hectare (stocking). For each of these outputs, residual values were plotted against predicted value, elevation, time increment (projection interval), and initial value.

No statistical tests are given in this analysis as repeated measurements have been taken from the basic experimental units (PSP). The consequences of this are:(i) estimators of

the regression coefficients may no longer have minimum variance but will still be unbiased and consistent; (ii) standard errors of coefficients in the regression will be underestimated; and, (iii) any significance tests or confidence limits constructed using *t* or *F* distributions are likely to be incorrect since assumed independence of errors is violated (West *et al*., 1984). The mean square error (MSE) for the regression is also likely to be underestimated if the correlation is positive and inflated if the correlations are negative (Snowdon *et al.,* 1999).

7.24 Calibration of 3PG

3-PG Data inputs

The 3-PG model (Landsberg and Waring, 1997) requires monthly average climatic inputs of solar radiation, mean air temperature, atmospheric vapour pressure deficit, rainfall, and frost days. If mean (T_a), maximum (T_x), and minimum (T_n) air temperatures are known, then T_a = ½(T_x + T_n). Vapor pressure deficit can also be estimated by the model from T_x and T_n, as half the difference between the saturated vapour pressure at T_x and T_n. The 3-PG model can be run using either actual monthly weather data or long term monthly averages. Other inputs are variables describing the physical properties of the site and initial biomass pools, latitude, a unit-less site fertility rating, maximum available soil water, minimum available soil water, initial available soil water, initial weight of foliage, initial weight of stems, foliage, and roots, and a general descriptor of the soil texture (Landsberg and Waring, 1997).

Initial Biomass Pool Inputs

Above ground initial biomass pools input into 3-PG were estimated using existing tree-level biomass equations for New Zealand (John Moore, 2005, pers. comm.) from the following equations;

$$\text{Stem mass} = 0.296054 + 0.01035918 \cdot DBH^2 \cdot H \qquad (7.9)$$

$$\text{Stem bark mass} = 0.00085 \cdot DBH^{1.906} \cdot H^{1.202} \qquad (7.10)$$

$$\text{Total foliage mass} = 0.07776 \cdot DBH^{2.858} \cdot H^{-1.410} \quad (7.11)$$

where: DBH = diameter at breast height (cm), H = height of the tree (m), and mass is in kg/tree.

Following Beets *et al*. (1999), initial root mass input into 3-PG was assumed to be 30% of total biomass.

Edaphic properties

Available soil water (ASW) levels and soil classification data were extracted from a local database of New Zealand soil types (Barringer *et al.* 1998) where soil attributes are linked to latitude and longitude. Maximum ASW, which was readily available in the database was used directly in 3-PG. Minimum ASW was also available in the database but was only used to estimate an initial value of ASW for the simulation which will be discussed further. Minimum ASW values input into 3-PG were set to 0 (mm) for all plots. This was done because inputting actual minimum values (which were greater than 0, see figure 7.2) would artificially prevent simulated trees from suffering significant stress during periods of drought and would tend to give incorrectly high estimates of LAI in model simulations (Landsberg and Waring, 1997).

Initial available soil water was estimated from values of maximum and minimum ASW available in the soils database acquired from Barringer *et al.* (1998). An estimate of initial ASW was made by a linear interpolation between values of maximum and minimum ASW depending on the initial month of the simulation (Figure 7.3). Months of maximum and minimum available soil water were based on measurements of root zone water balance in Canterbury detailed by Richardson *et al.* (2002) (Figure 7.2).

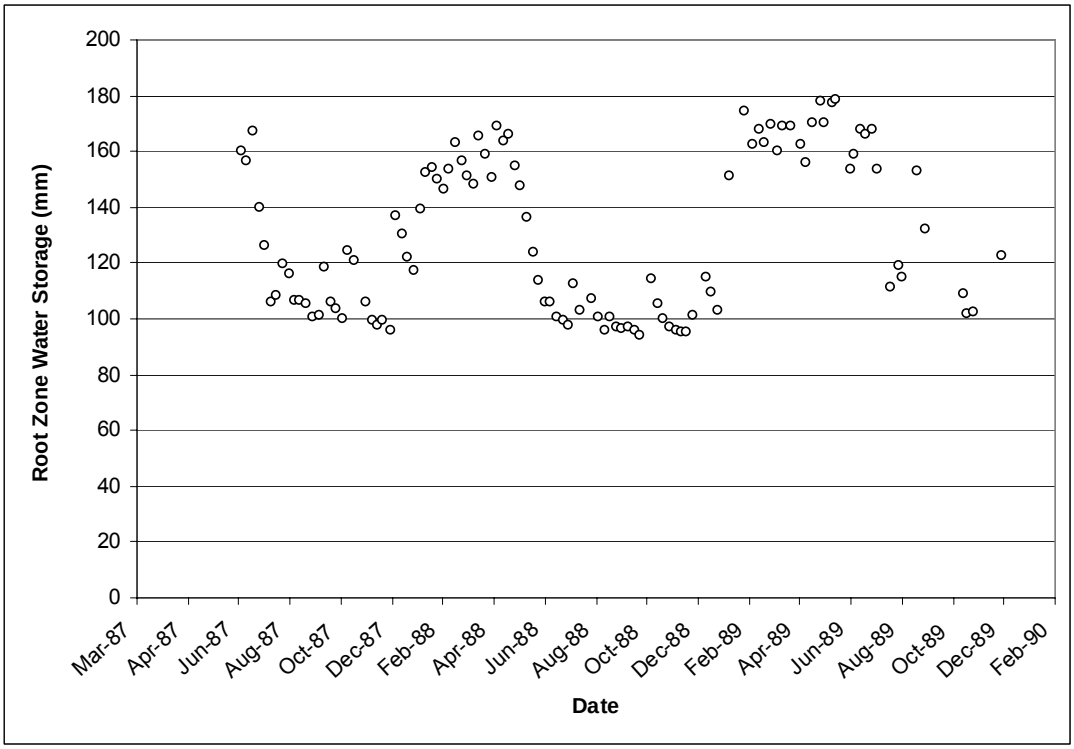


Figure 7.2: Root zone water content (mm) measured in Burnham New Zealand between March 1987 and February 1990. Reproduced from Richardson *et al.* (2002).

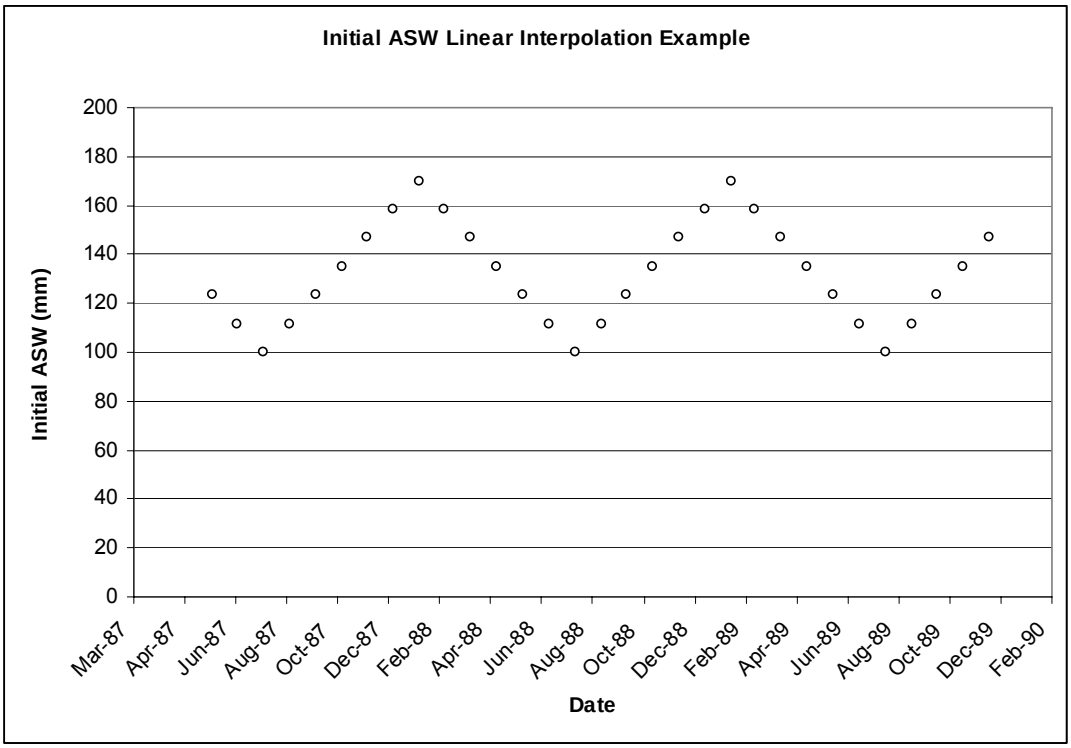


Figure 7.3: Estimates for initial ASW for a single plot where for this illustration the maximum and minimum values are assumed to be 170 and 100 mm respectively.

Soil classes were translated from the soils database format in Barringer *et al.* (1998) to 3-PG soil classes by referring to the soil texture triangle in McLaren and Cameron (1996). The conversion of soil classes from the database format to 3PG soil classes is listed in Table 7.2. Site fertility rating is a unit-less input into 3PG, ranging from 0 to 1, that is used to characterise the soil fertility of a given site. This input was fit to each of three forest types which will be discussed further in the parameterisation section. Final values of site fertility rating by forest type are listed in Table 7.3.

Table 7.2: Soil types used in 3-PG.

Baringer *et al.* (1998) soil class	3-PG soil class
Complex	Clay Loam
Hill soils	Clay Loam
Sand	Sandy
Sandy loam	Sandy Loam
Shallow sandy loam	Sandy Loam
Shallow silt loam	Clay Loam
Shallow soils	Sandy
Silt loam	Clay Loam
Stony and shallow silt loam	Clay Loam
Stony and very stony silt loam	Clay Loam
Stony loam and sandy loam	Sandy Loam
Stony sandy loam	Sandy Loam
Very stony sand and very stony sandy loam	Sandy Loam
Very stony sandy loam	Sandy Loam

Stand state inputs

Initial stand state variables of latitude, year planted, month planted, initial year, initial month, end age, and initial stocking are available from the database used for model validation which was previously discussed.

Climatic inputs

The climatic inputs of solar radiation, minimum air temperature, maximum air temperature, and rainfall were estimated for each site using the method of correcting ten year average values for each site by local met station measurements. Actual inputs are based on a system of corrected estimates from BIOCLIM (Leathwick *et al.* 1998) climate

surfaces. BIOCLIM is a set of surface equations for New Zealand that estimate climatic variables on a ten year average based on interpolation between climate station measurements (Leathwick *et al.* 1998). BIOCLIM outputs monthly averages of temperature, wind, rainfall, radiation, and average humidity extremes across New Zealand based on location. To simulate the monthly variation in actual climate readings, as opposed to mean monthly values over a 10-year period, the long-term mean monthly climate values derived from BIOCLIM were rescaled using monthly climate measurements from a nearby weather station. These estimates were corrected by reference point measurements made at Christchurch Airport (S43.5°, E172.55°, elevation 37m). Differences between climatic estimates from BIOCLIM and measurements at Christchurch Airport were calculated for each month included in the study from 1984 – 2004. Corrections were added to long term monthly estimates at each permanent sample plot. So instead of using the same monthly averages every year, a new value was calculated every month based on actual climate at Christchurch Airport to reflect years of extreme climatic events such as droughts. Similar approaches of correcting long term averages with local measurements have been used by Tickle *et al.* (2001), and Snowdon *et al.* (1999). This process is described in detail in Chapter 6.

Parameterisation

Parameterisation of the model was achieved by (1) adjusting selected parameters to obtain good fits of model output to observed measurements of stand basal area and stocking on a subset of calibration plots and (2) through the examination of time series estimates of LAI from the 3-PG model against LAI estimates produced by model 1 presented in chapter 5. The calibration data set was comprised of 1059 plot measurements taken within 200 plots that was representative of all elevation ranges, age classes, and re-measurement intervals within the original data set for modelling described in chapter 4. This method of parameterisation is similar to that used by Sands and Landsberg (2002), for stands of plantation-grown *Eucalyptus globulus* in Tasmania and Western Australia. A final set of parameter inputs to the model are listed in table 7.3, indicating whether the value used in the simulation was obtained directly from observed data (observed), was estimated by fitting output from the model to observed data (fitted), or was some generic model default (default).

Table 7.3: Description and source of 3-PG parameters for *Pinus radiata* in Canterbury New Zealand.

Meaning/comments	Parameter Source	*P.radiata* (mine)	Units
Allometric Relationships and Partitioning			
Ratio of foliage: stem partitioning at D = 2 cm	Fitted	Table 7.33	-
Ratio of foliage: stem partitioning at D = 20 cm	Fitted	Table 7.33	-
Constant in stem mass and diameter relationship	Observed	0.0095	-
Power in stem mass and diameter relationship	Observed	2.9352	-
Maximum fraction of NPP to roots	Default	0.8	-
Minimum fraction of NPP to roots	Default	0.25	-
Temperature and frost modifier (fT)			
Minimum temperature for growth	P. radiata Default	0	deg. C
Optimum temperature for growth	P. radiata Default	20	deg. C
Maximum temperature for growth	P. radiata Default	32	deg. C
Number of days of production lost for each frost day	P. radiata Default	1	Days
Soil water modifier (fSW)			
Moisture ratio deficit which gives fq = 0.5	P. radiata Default	0.7	-
Power of moisture ratio deficit in fq	P. radiata Default	9	-
Fertitlity effects			
Value of 'm' when FR = 0	P. radiata Default	0	-
Value of 'fNutr' when FR = 0	P. radiata Default	0.6	-
Power of (1-FR) in 'fNutr'	P. radiata Default	1	-
Age modifier (fAge)			
Maximum stand age used to define relative age	P. radiata Default	50	Years
Power of relative age in function for fAge	P. radiata Default	4	-
Relative age to give fAge = 0.5	P. radiata Default	0.5	-
Litterfall & root turnover			
Maximum litterfall rate	Raison *et al. (1992)*	0.025	1/month
Litterfall rate for very young stands	P. radiata Default	0.001	1/month
Age at which litterfall rate has median value	Raison *et al. (1992)*	36	Month
Average monthly root turnover rate	P. radiata Default	0.015	1/month
Conductance			
Maximum canopy conductance	P. radiata Default	0.02	m/s
LAI for maximum canopy conductance	P. radiata Default	3.33	-
Defines stomatal response to VPD	P. radiata Default	0.05	1/mBar
Canopy boundary layer conductance	P. radiata Default	0.2	m/s
Stem numbers			
Max. stem mass per tree for 1000 trees/hectare	Fitted	Table 7.33	kg/tree
Power in self-thinning rule	P. radiata Default	1.5	-
Fraction mean single-tree foliage	P. radiata Default		
Biomass lost per dead tree		1	-
Fraction mean single-tree root biomass lost per dead tree	P. radiata Default	0.2	-
Fraction mean single-tree stem biomass lost per dead tree	P. radiata Default	1	-

Canopy structure and processes			
Specific leaf area at stand age 0	P. radiata Default	5	m2/kg
Specific leaf area for mature leaves	P. radiata Default	5	m2/kg
Age at which specific leaf area = (SLA0+SLA1)/2	P. radiata Default	2	Years
Extinction coefficient for absorption of PAR by canopy	P. radiata Default	0.5	-
Age at canopy cover	P. radiata Default	3	Years
Maximum proportion of rainfall evaporated from Canopy	P. radiata Default	0.15	-
LAI for maximum rainfall interception	P. radiata Default	5	-
Canopy quantum efficiency	Fitted	0.046	molC/molPAR
Ratio NPP/GPP	P. radiata Default	0.47	-
Branch and bark fraction (fracBB)			
Branch and bark fraction at age 0	P. radiata Default	0.5	-
Branch and bark fraction for mature stands	P. radiata Default	0.1	-
Age at which fracBB = (fracBB0+fracBB1)/2	P. radiata Default	5	Years
Basic Density			
Minimum basic density - for young trees	Cown *et al*. 1991	0.360	t/m3
Maximum basic density - for older trees	Cown *et al*. 1991	0.463	t/m3
Age at which rho = (rhoMin+rhoMax)/2	Cown *et al*. 1991	11	Years
Conversion factors			
Intercept of net and solar radiation relationship	P. radiata Default	-90	W/m2
Slope of net and solar radiation relationship	P. radiata Default	0.8	-
Molecular weight of dry matter	P. radiata Default	24	gDM/mol
Conversion of solar radiation to PAR	P. radiata Default	2.3	mol/MJ

Rationale behind fitted and observed parameters

The fitted parameters were the ratio of foliage: stem partitioning at D = 2 cm (pFS2), ratio of foliage: stem partitioning at D = 20 cm (pFS20), maximum stem mass per tree for 1000 trees/hectare (wSx1000), canopy quantum efficiency (gamma), and site fertility factor.

The parameter affecting stem numbers per hectare (wSx1000), was varied by forest type (sands, plains, or hills forests) to reduce the difference between predicted and observed basal area, stocking, and leaf area (where observed leaf area was assumed to be LAI predicted by model 1 or equation 5.3 from chapter 5). Table 7.4, shows the final fitted values for wSx1000 by forest type.

The parameter for canopy quantum efficiency (gamma) was kept constant for all forest types, and a final value was fitted in the same manner as that used to fit wSx1000. A

single value was decided upon as it was assumed that canopy quantum efficiency would not be responsible for differences in growth between forest types as all contained the same species. Rather, it was assumed that site differences in soil water, soil texture, and climate should account for differences in growth. A final fitted value for gamma is listed in Table 7.3.

Allometric relationships and partitioning ratios pFS2 and pFS20 were varied by forest type (sands, plains, or hills forests) to reduce the difference between predicted and observed basal area, stocking, and leaf area. The general method used to vary these coefficients was by examining trends in residual plots and adjusting pFS2 and pFS20 so that the biomass partitioning coefficients of ηf , and ηs in 3PG would be reduced or increased to change the allocation of biomass to foliage and stems in a way that reduced residual errors. An example of this relationship for coastal sands forests where pFS2 = 0.7, and pFS20 = 0.4 is shown in Figure 7.4 . Equations for the relationship between pFS2, pFS20, ηf , and ηs are derived in Sands and Landsberg (2002). The final parameterised values of pFS2 and pFS20 by forest type are listed in Table 7.4.

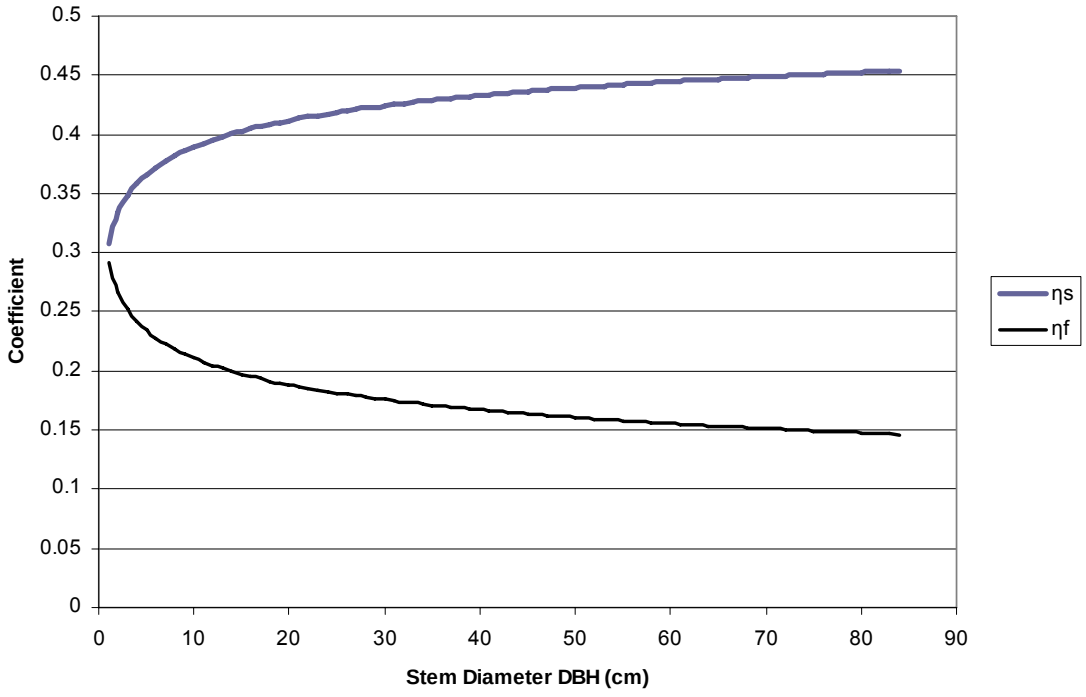


Figure 7.4: Biomass and Partitioning Coefficients for stems (ηs) and foliage (ηf) in 3PG for coastal sands forest type. Assuming pFS2 = 0.7, pFS20 = 0.4.

Table 7.4: Fitted values of 3PG parameters by forest type.

Forest Type	**3PG parameter**			
	pFS2	pFS20	wSx1000	Site Fertility Rating
Sands	0.7	0.4	150	1
Plains	0.65	0.4	150	.75
Hills	0.8	0.2	250	1

Basic density values were obtained from Cown (1992), and Cown *et al.* (1991). Minimum basic density for young trees in Canterbury was 350 (Kg / m^3), while maximum basic density for old trees in Canterbury was 450 (Kg / m^3). The age at which average density = (Min density + Max density)/ 2 occurred was 11 years (Cown *et al.* 1991). Since these values of density were for extracted wood they were corrected to account for the percentage of resin content in un-extracted wood. Following Cown (1992), an average value of 3 percent of the extracted density was added to account for resin. Final values for these parameters input into 3PG are listed in Table 7.3.

Maximum litterfall rate and age at which litterfall rate has a median value corresponded well with results presented in Raison *et al.* (1992). The parameter used for maximum litterfall rate was 0.025 (1/month) which corresponds to needles being retained on trees for about 3.3 years. This value corresponds with Raison *et al.* (1992) who reported an average needle life of 2-4 years for radiata pine depending on water stress conditions. The default value of age at which litterfall rate has median value (3 years) also corresponded well with results presented in Raison *et al.* (1992).

The allometric relationship parameters of the constant and power terms in the stem mass and diameter relationship (Stem Const, and StemPower) were observed. Measurements of stem diameter were modelled against mass per tree using an exponential equation form to solve for the constant and power term in this relationship (Figure 7.5). Tree mass was determined as the product of density (which was assumed to be 400 (kg m^{-3}) from Cown *et al.* (1991)) and volume. Volume was determined from tree height and diameter using an equation developed for the Canterbury plains by Zhao (1999).

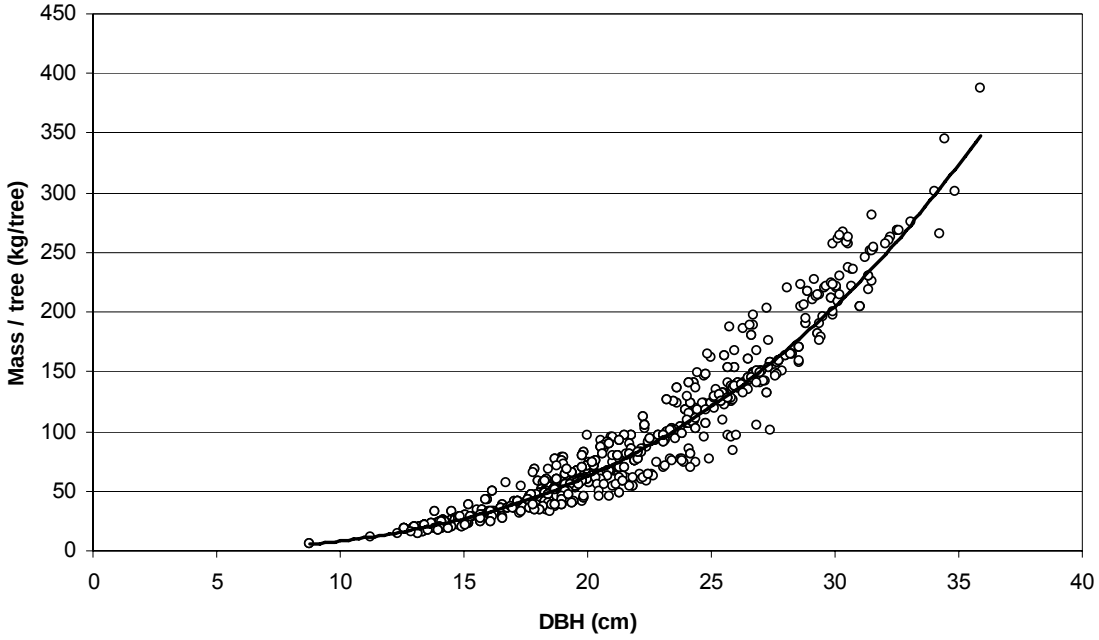


Figure 7.5: Relationship between the mass per tree and DBH. The fitted line was used to find the constant and power terms for the allometric parameters of Stem Const, and StemPower. $y = 0.0095 \cdot x^{2.9352}$ Where: y = mass / tree (kg) and x = DBH (cm).

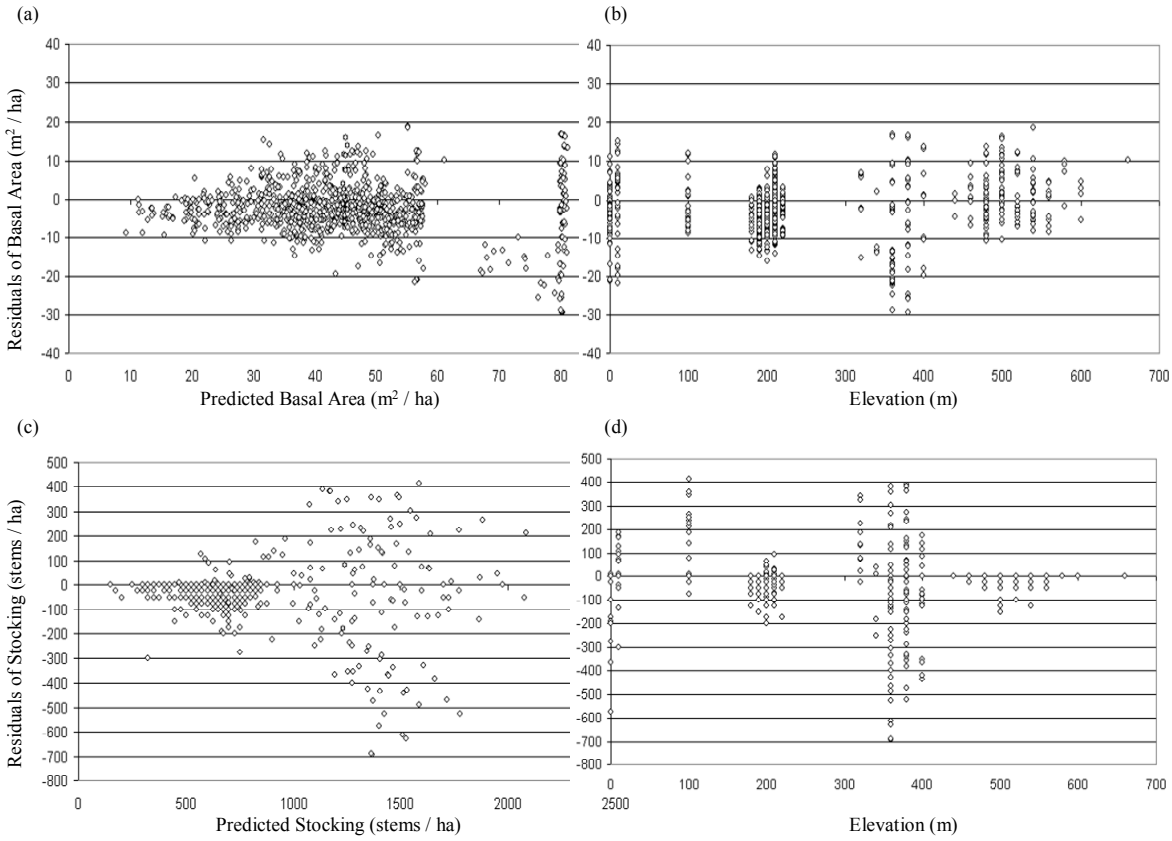


Figure 7.6: Calibration residual plots for 3PG showing (a) G residuals versus predicted, (b) G residuals versus elevation, (c) stocking residuals versus predicted, and (d) stocking residuals versus elevation.

7.3 Results

CANTY Validation

Model components of mean top height, basal area, and final stocking were validated by the display of graphical residual plots. Plots display residual values against predicted, elevation, interval length, and initial value (Figures 7.7, 7.8, and 7.9). Fitting statistics of mean square error (MSE), average model bias (AMB), and model efficiency factor (EF) were also calculated, and are listed in Table 7.5.

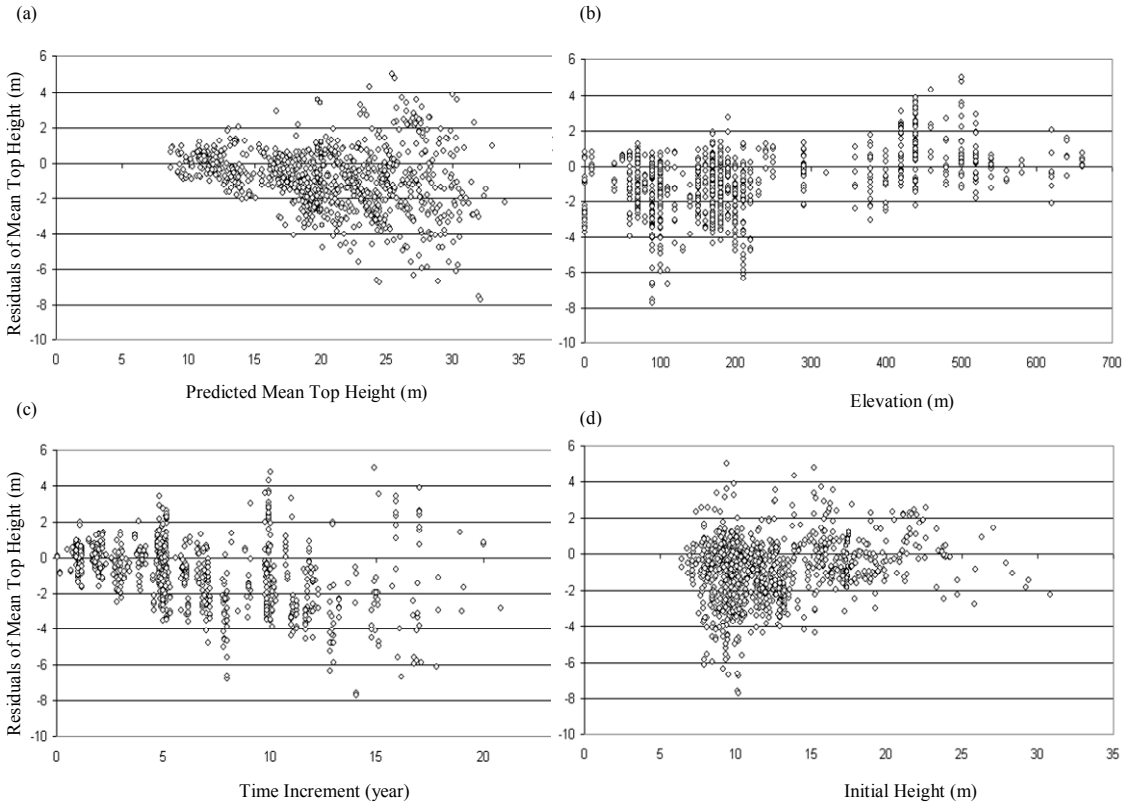


Figure 7.7: Mean top height residual plots for CANTY validation showing (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial height.

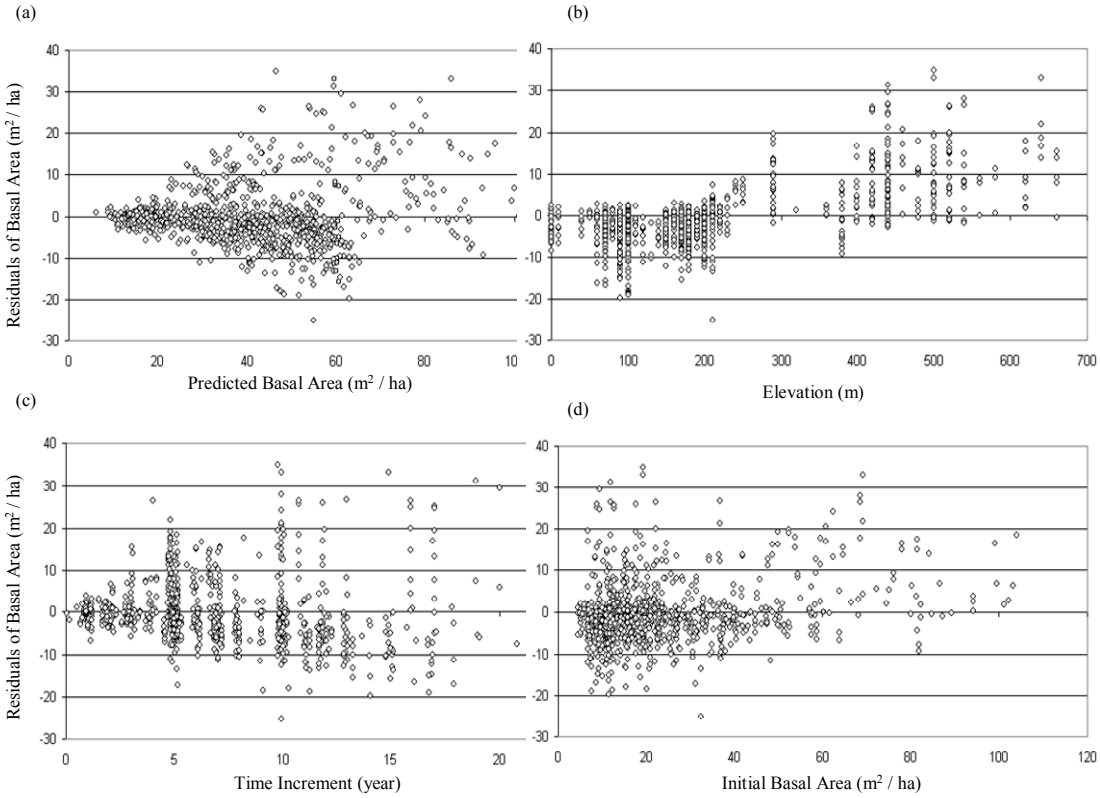


Figure 7.8: Basal area residual plots for CANTY validation showing (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial basal area.

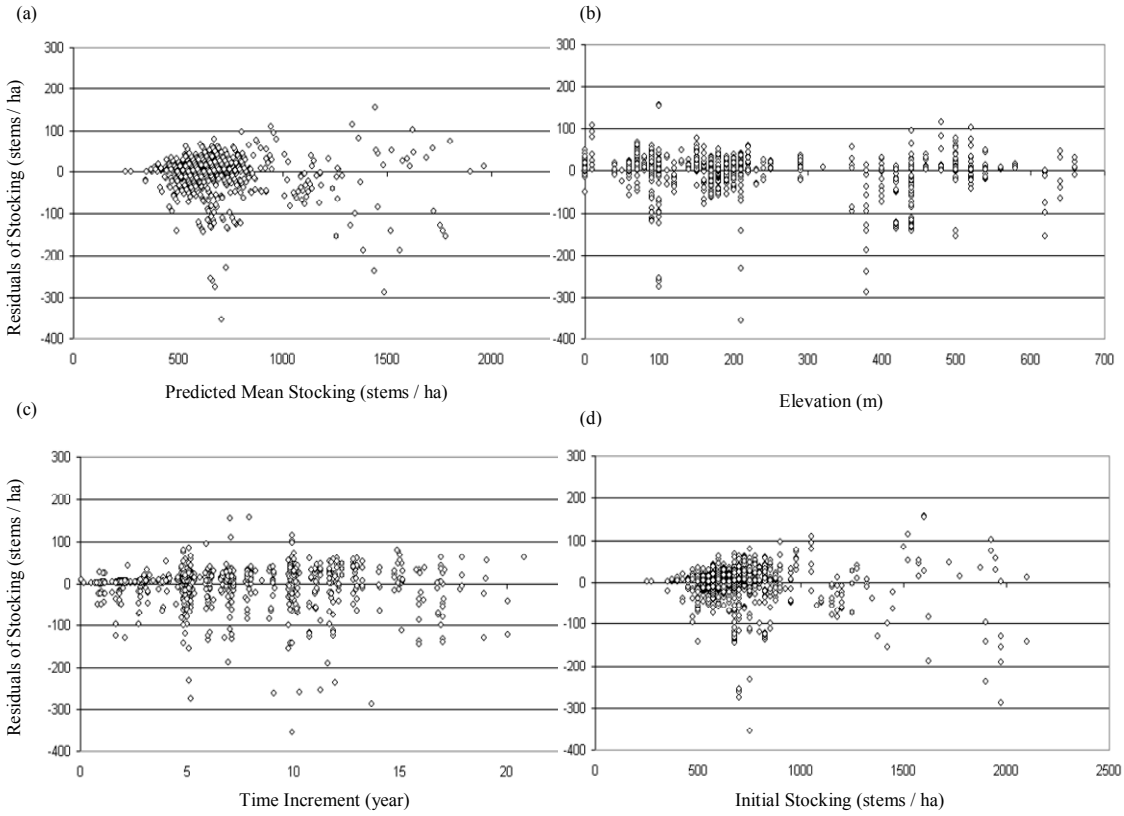


Figure 7.9: Stocking residual plots for CANTY validation showing (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial stocking.

CANTY showed severe bias in both mean top height and basal area against elevation. These effects were most apparent in the plots against elevation and re-measurement interval (Figure 7.7 b and c, and Figure 7.8 b and c). However, residual plots of stocking show little bias for predicted value, elevation, time increment, or initial value (Figure 7.9 a-d).The model CANTY was originally fit with data at lower elevations and with data of shorter average intervals than the data used for validation in this study. Because state space models are critically sensitive to site index the severe bias in mean top height and basal area plots may have been due to the fact that that the model was fit with short interval data from plains elevations only.

3PG Validation

Model components of basal area, and final stocking were validated by the display of graphical residual plots. Plots display residual values against predicted, elevation, interval

length, and initial value (Figures 7.10, and 7.11). Fitting statistics of mean square error (MSE), average model bias (AMB), and model efficiency factor (EF) were also calculated, and are listed in Table 7.5.

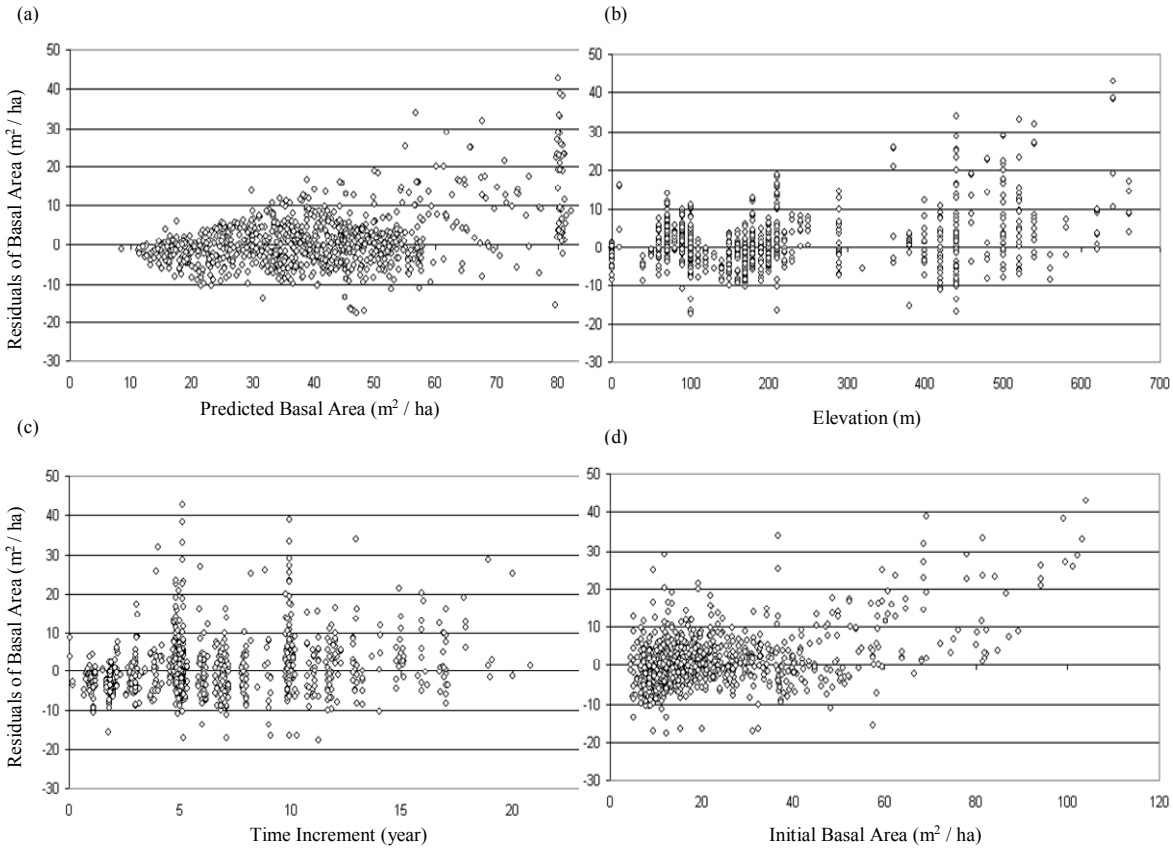


Figure 7.10: Basal area residuals plots of 3PG validation: (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial basal area.

Residual plots indicate bias for G projection at higher elevation plots (Figure 7.10 a), that was not apparent in the calibration of the model (Figure 7.6 b). This bias may have been due the effect of initial basal area on final projection of basal area (Figure 7.10 d). Basal area bias may also indicate an inability for 3PG to deal with calibration on such a large scale without fine tuning all parameters for fitting such as the photosynthetic efficiency (alpha). Fitting the parameter alpha for each of the three elevation classes may have improved the calibration and validation fit. This parameter was not fit for the three forest types as they all contain the same species which was assumed to have the most influence over photosynthetic efficiency.

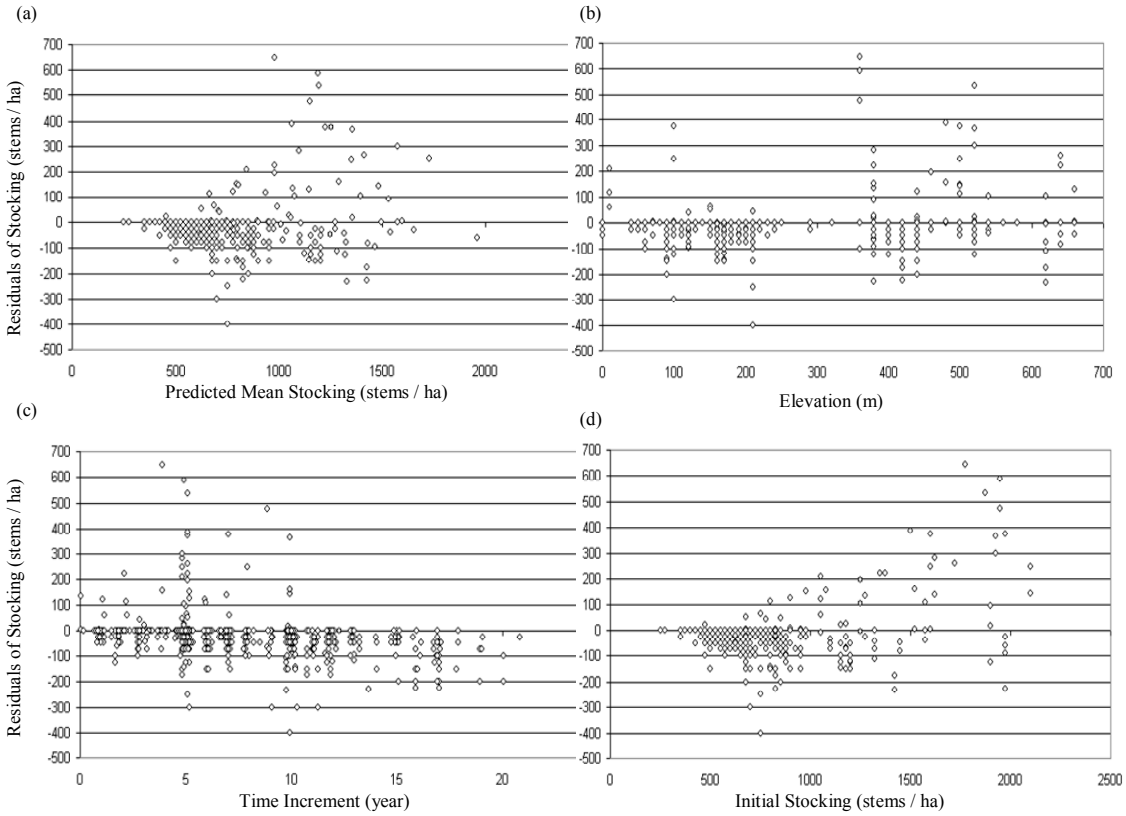


Figure 7.11: Stocking residual plots of 3PG validation: (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial stocking.

3 PG also showed an inability to deal with the prediction of stocking at the end of the simulation. This may have been due to mortality events not associated with overcrowding, as 3PG uses the -3/2 power law alone to estimate mortality. Bias in mortality prediction is apparent in both the calibration residual plots (Figure 7.6 c and d) and validation plots (Figure 7.11 a - d), although the bias is more severe in the validation.

CanSPBL(1.2) Validation Residual Plots

Model components of mean top height, basal area, and final stocking were validated by the display of graphical residual plots. Plots display residual values against predicted, elevation, interval length, and initial value (Figures 7.12, 7.13, and 7.14). Fitting statistics

of mean square error (MSE), average model bias (AMB), and model efficiency factor (EF) were also calculated, and are listed in Table 7.5.

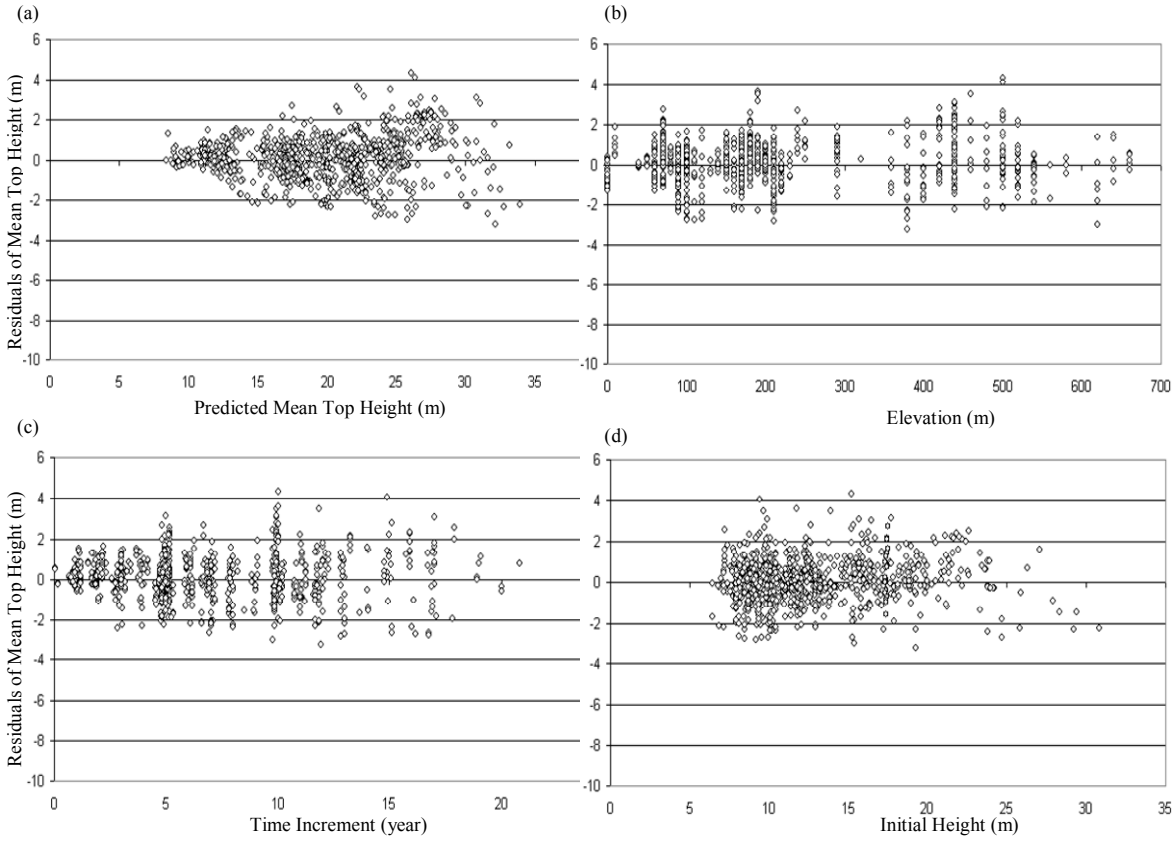


Figure 7.12: Mean top height residuals plots of CanSPBL(1.2) validation: (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial height.

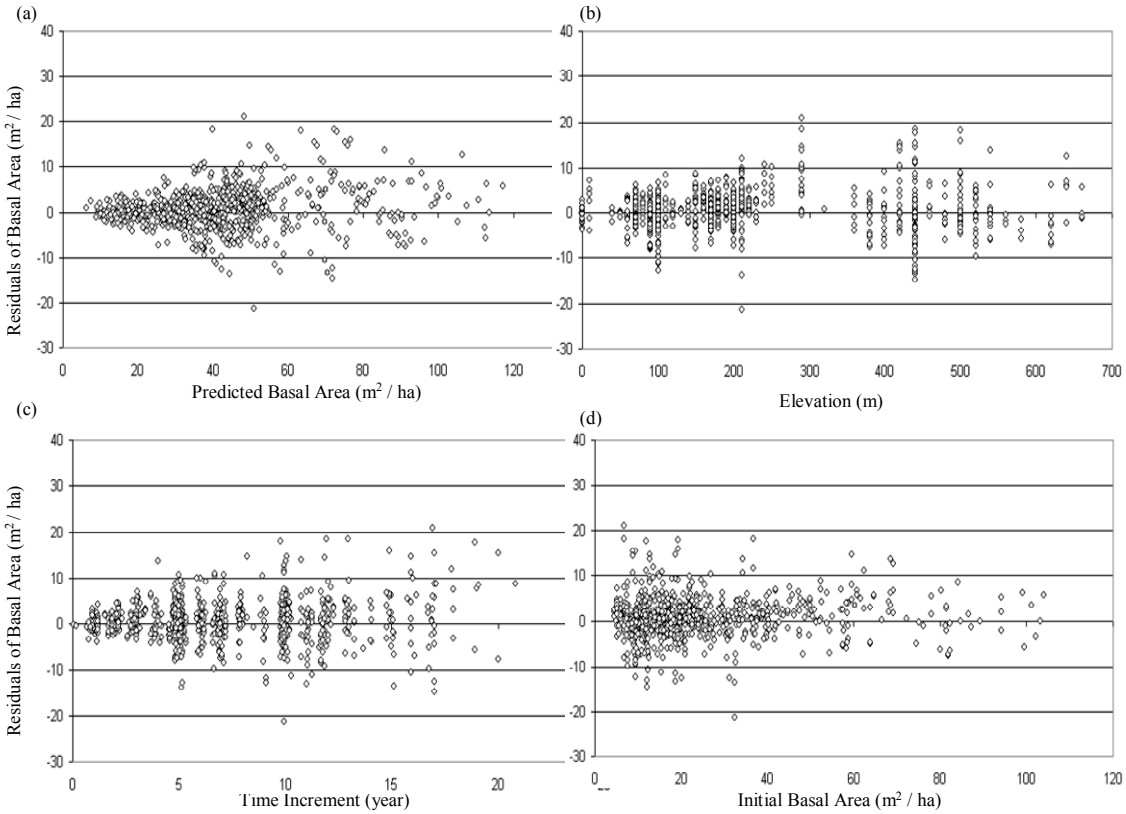


Figure 7.13: Basal area residuals plots of CanSPBL(1.2) validation: (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial basal area.

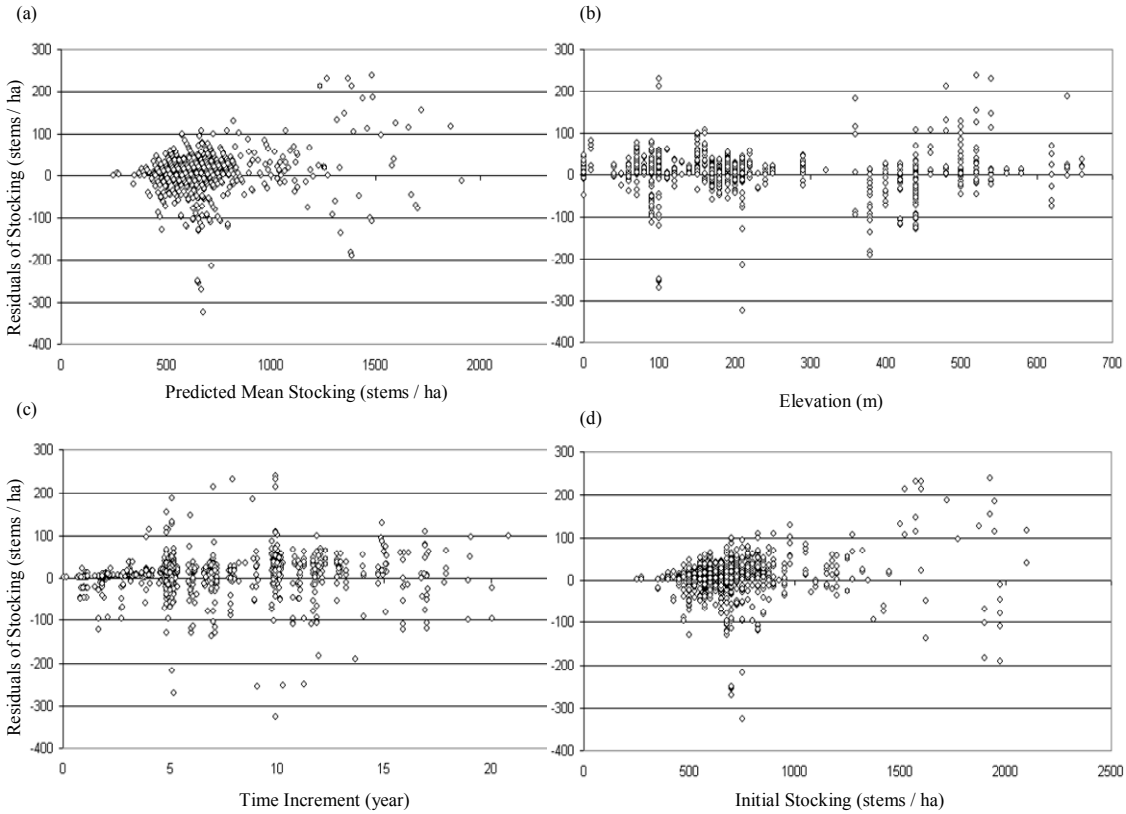


Figure 7.14: Stocking residual plots of CanSPBL(1.2) validation: (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial stocking.

Residual plots of mean top height and basal area showed little bias. The basal area model indicated possible bias in under prediction between elevations of 250 to 350 m (Figure 7.13 b). Stocking residuals indicate bias at plots above 450 m (Figure 7.14 b), where the model is underestimating stocking or at higher elevations. Stocking bias may have been an artefact of the way the mortality model was fitted in CanSPBL (1.2) to make predictions that seemed reasonable to managers. This process involved segregating the dataset to only include non-catastrophic mortality events so that mortality would not be overestimated on average plots. The validation dataset was not segregated in this way and residual bias at higher elevations may be due to catastrophic mortality events.

CanSPBL(water) Validation Residual Plots

Model components of mean top height, basal area, and final stocking were validated by the display of graphical residual plots. Plots display residual values against predicted,

elevation, interval length, and initial value (Figures 7.15, 7.16, and 7.17). Fitting statistics of mean square error (MSE), average model bias (AMB), and model efficiency factor (EF) were also calculated, and are listed in Table 7.5.

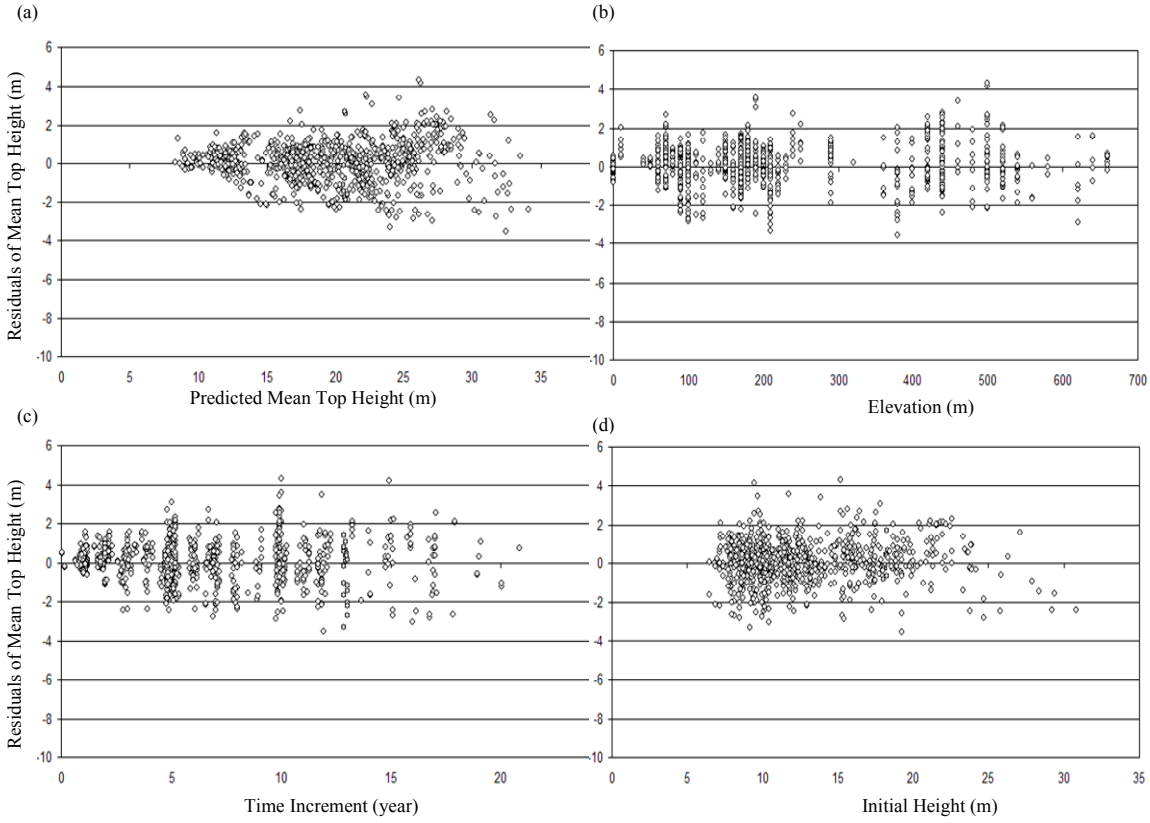


Figure 7.15: Mean top height residuals plots of CanSPBL(water) validation: (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial height.

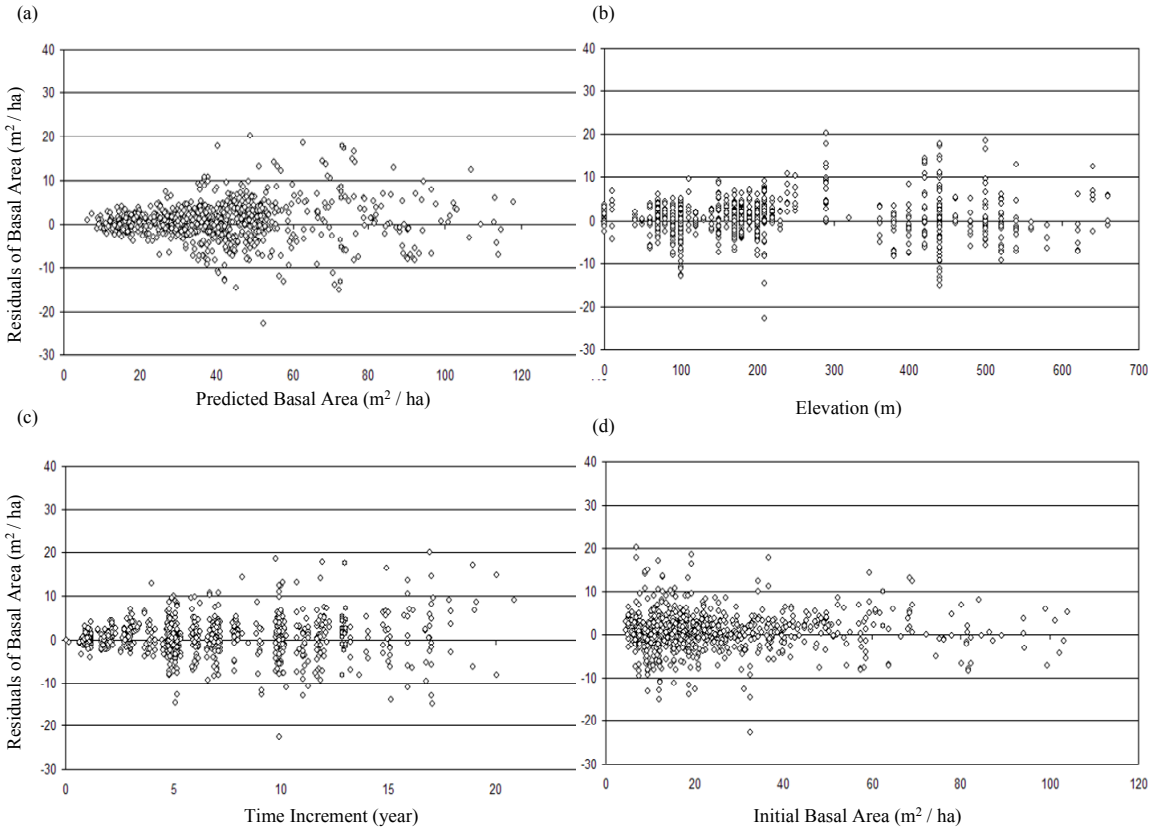


Figure 7.16: Basal area residuals plots of CanSPBL(water) validation: (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial basal area.

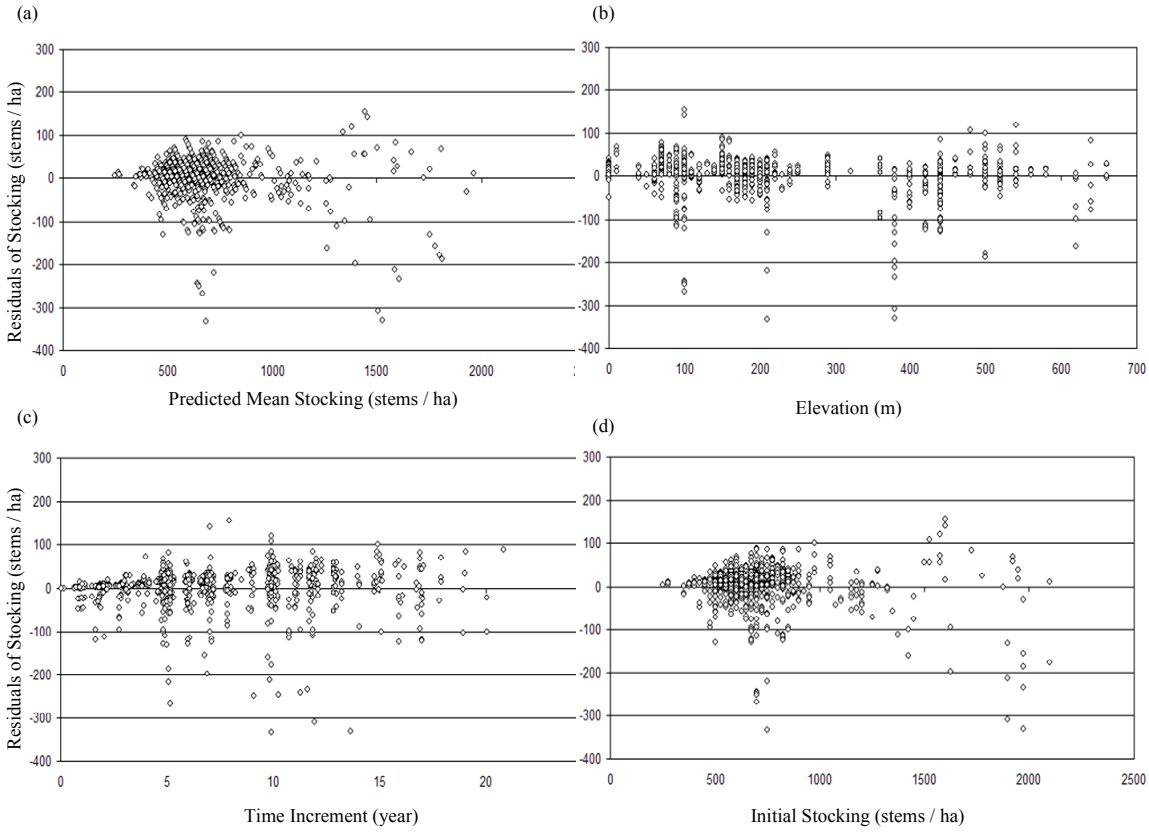


Figure 7.17: Stocking residual plots of CanSPBL(water) validation: (a) residuals versus predicted, (b) residuals versus elevation, (c) residuals versus time increment, and (d) residuals versus initial stocking.

Residual plots of mean top height and basal area showed little bias. The basal area model indicated possible bias in under prediction between elevations of 250 to 350 (m) (Figure 7.16 b). Stocking residuals indicate bias for plots located above 450 m (Figure 7.17 b). Stocking bias may have also been an artefact of the way the mortality model was fitted in CanSPBL (water) to make predictions that seemed reasonable to managers. This process involved segregating the dataset to only include non-catastrophic mortality events so that mortality would not be overestimated on average plots. The validation dataset was not segregated in this way and residual bias at higher elevations may be due to catastrophic mortality events.

Comparison of Models

Model fitting statistics for each tested model component are listed in terms of mean square error, average model bias, and model efficiency in Table 7.5. Model residual distribution statistics for each tested model component in terms of skewness, kurtosis, and p-value for the test of normality of residuals are listed in Table 7.6.

Table 7.5: Statistics of model fit for validation datasets against CanSPBL(1.2), CANTY, 3PG, and CanSPBL(water) in terms of Mean Square Error (MSE), Average Model Bias (AMB), and Model Efficiency Factor (EF).

Model		**Statistics of Model Fit**		
		MSE	AMB	EF
CanSPBL(1.2)				
	Stocking	2161.33	6.51	0.96
	Basal Area	17.99	0.55	0.96
	Mean Top Height	1.11	0.08	0.96
CANTY	Stocking	2022.45	-3.17	0.96
	Basal Area	55.00	-0.22	0.87
	Mean Top Height	3.81	-0.91	0.87
3PG	Stocking	5566.82	-17.94	0.89
	Basal Area	53.48	1.08	0.88
CanSPBL(water)				
	Stocking	2100.67	1.94	0.96
	Basal Area	17.32	0.46	0.96
	Mean Top Height	1.07	0.06	0.96

Table 7.6: Model residual distribution statistics for CanSPBL(1.2), CANTY, 3PG, and CanSPBL(water) in terms of Skewness, Kurtosis, and p-value for the Kolmogorov-Smirnov test for normality of residuals.

Model		**Residual Distribution Statistics**		
		Skewness	Kurtosis	p-value
CanSPBL(1.2)				
	Stocking	-1.01	11.01	0.01
	Basal Area	0.42	3.72	0.01
	Mean Top Height	0.13	0.95	0.01
CANTY	Stocking	-2.59	12.23	0.01
	Basal Area	-2.52	23.90	0.01
	Mean Top Height	-0.29	0.95	0.01
3PG	Stocking	2.54	23.84	0.01
	Basal Area	1.61	5.07	0.01
CanSPBL(water)				
	Stocking	-2.79	13.97	0.01
	Basal Area	0.22	4.04	0.01
	Mean Top Height	-0.02	1.04	0.01

Overall, the models CanSPBL(water), and CanSPBL(1.2) performed the best in terms of basal area and mean top height prediction. Both models CanSPBL(water), and CanSPBL(1.2) showed a slightly worse fit in predictions of stocking than did the model CANTY (Table 7.5). The hybrid model 3PG showed a better fit for the prediction of basal area than the statistically based model CANTY, but showed a worse fit for the prediction of final stocking than all other models (Table 7.5).

In terms of distribution of residuals, CanSPBL(1.2) and CanSPBL(water) performed best in terms of skewness. On average kurtosis was lowest for CanSPBL(1.2), followed by CanSPBL(water). 3PG performed the worst on average, in terms of the distribution of residuals, and all models tested positively for the normality of residual distribution (Table 7.6).

7.4 Discussion

This chapter illustrates the utility of process-based modelling. Interesting results of this study show that the process-based model 3PG predicted basal area more precisely than a statistically based (state-space) model CANTY even though stem numbers predictions were less precise. This is an interesting result as it shows either CANTY's inability to deal with changing growing conditions at different elevations, or it shows possible changes in growing environment from the time of initial model fitting such as nutrient losses in the soil. 3PG also predicted basal areas that were comparable to those predicted by the models CanSPBL(water), and CanSPBL(1.2). This result shows that process-based models can give comparable results for basal area prediction as statistically-based models while maintaining the ability to make predictions under changing environmental conditions, such as varying climate and soil type. The model 3PG also has many physiological outputs such as estimated carbon pools, and leaf area that can be of use to forest managers interested in carbon accounting under the Kyoto protocol. The model did not perform as well for the prediction of stocking compared to the other models tested. This effect may have been due to 3PG's use solely of the -3/2 power law to predict mortality events. The -3/2 power law would only predict mortality events due to overcrowding and would be of little use in cases of mortality caused by wind, or insect attacks. Further, basal areas output from 3-PG are determined by inverting the allometric equation that describes the relationship between individual stem mass and diameter, calculating average basal area per tree and then multiplying by stem number, so if predicted stem numbers are inaccurate, then basal area will also be affected.

The state-space model CANTY performed the worst in terms of basal area and mean top height prediction. This may have been a result of the critical dependence of state space models on site index and that the height model is used instead of time when predicting basal area and stocking. If the height model is biased then the others are also likely to be biased. As the model CANTY was fitted to data of only short re-measurement intervals and lower elevation sites on the Canterbury plains, the model did not perform well predicting mean top height at elevations and at longer intervals. Site index is calculated as mean top height of the stand at age 20, so the problem with fitting mean top height can

be compounded when trying to predict basal area with site indices that were originally biased.

CHAPTER 8

General Discussion

A broad research area was covered in this study, beginning with the validation of an existing growth and yield model, creating a new growth and yield model with a more complete data set, building a sub-model to predict canopy leaf area index, incorporating an index for root zone water balance into the updated growth and yield model, and then comparing the accuracy of a range of models from empirical growth and yield models to physiological process-based models via a residual analysis. This chapter will discuss new features in this study, some relevant specific points of the research and future research in modelling.

8.1 New Features in This Research

8.1.1 Improvements in the model CanSPBL(1.0) in comparison with the model CanSPBL(1.2)

The existing model CanSPBL(1.0) is a stand level model of radiata pine plantations growing in the Canterbury region of New Zealand within the Selwyn Plantation Board estate. The model outputs 5 stand state variables that describe mean top height, basal area per hectare, stems per hectare, and diameter distribution. Sites on the Selwyn estate vary with elevation that cover 3 forest types of plains, hills, and coastal sands with elevations ranging from 2 – 600 metres above sea level. Distinctions between these forest types are due to both elevation and soil type. With increasing distance from the sea, soils underlying forests in Canterbury change from coastal sands, shallow and dry floodplains soils, to deep wet LOESS hill soils (Barringer *et al.,* 1998). Average temperatures are 10.9 degrees for plains and coastal forests, and 11.6 degrees for forests growing on the hills. Annual rainfall increases from 600 to 1100mm for coastal-plains and hills forest respectively (Leathwick *et al.* 1998). The model CanSPBL(1.0) was validated with an independent dataset that indicated bias in prediction of forest growth at higher elevation sites. The original model was not intended to distinguish between the three distinct forest types of sands, plains and hills. The original model CanSPBL(1.0) did include the

predictor variable of elevation, and data from higher elevation sites was limited in model building. This lack of distinction resulted in bias in stand level projections.

An updated model CanSPBL(1.2) was created. This model encompasses a larger state space and varies predictions of growth and stand volume according forest type and elevation. Refitting the equations that predict basal area, mean top height, standard deviation of diameter, and stocking in the growth and yield model CanSPBL(1.0) with a more current data set that better covered a range of elevations and stand ages on the Selwyn estate showed a significant improvement in model accuracy. The effects of elevation incorporated into this model were somewhat complicated and indicated elevation may be a proxy for physiological factors that effect growth or variations in site factors. Seinbrenner (1963), and Grey (1979) demonstrated how soil factors had a highly significant influence on the growth of Douglas fir and radiata pine. In New Zealand, Ballard (1971) found a strong correlation between an index for growing potential (site index) and slope, aspect, drainage, mycorrhizal associations, soil pH, and soil phosphates. Climate has also been shown to affect the growth of radiata pine by van Laar (1967).

8.1.2 Modelling canopy Leaf Area Index

Two final models were developed to predict stand-level LAI from stand state variables. LAI estimated with these models can be used as an input in ecosystem models that simulate carbon balance and hydrological cycles. The two final models were surprisingly simple in form, depending only on stand age, mean top height, and elevation. Results from this study demonstrate the insensitivity of LAI to stocking in older stands. During model fitting, there were no apparent patterns in residual plots against other possible predictor variables, such as stocking and rainfall. This pattern of simplicity has also been reported by Pinkard and Neilsen (2003), who found similar results for stand leaf areas in plantation-grown eucalypts. They reported canopy leaf areas that were nearly independent of plantation spacing, having only a moderate effect at spacings below 833 trees/ha. These results are also supported by Savill *et al.* (1997) who summarised the general effects of spacing in even aged stands as: once canopy closure has occurred and a site is being fully used in widely spaced stands, dry matter increment is similar to that in close spacings on the same site.

8.1.3 Improvements in the model CanSPBL(1.2) and creation of CanSPBL(water)

An updated model CanSPBL(water) was created. This model incorporates and index of root zone water balance over the simulation period and varies predictions of growth and stand volume according forest type and elevation. Refitting the equations that predict basal area, mean top height, maximum diameter, standard deviation of diameter, and stocking in the growth and yield model CanSPBL(1.2) showed a small but significant improvement in model accuracy. The effects of elevation incorporated into this model were simpler than those incorporated into the model CanSPBL(1.2). The updated version of the model included both ASW deficit and a simplified form of elevation. This result indicates that the effect of elevation in the Canterbury region describes water relations on the site, and may also be a proxy for other site factors such as nutrient availability, or possibly average temperature. A model that directly incorporates water balance, however, is likely to be more robust that one that relies on proxies like elevation.

Precision increased slightly with interval length as shown in residual plots (Figures 6.4, 6.5, 6.6, and particularly Figure 6.8). This effect may have been due to the levelling of water balance models precision increasing with time of simulation. If initial water balance estimates are different from measurements, they become more precise over time for months following initialization, and this may affect growth and yield model precision for longer measurement intervals.

Validation with average climatic inputs increased precision of the models that predict stocking, basal area, and maximum diameter. Predicted stocking showed the largest reduction of residual mean square error from the original updated version (CanSPBL(1.2), at 7.81% using average climate. Including causal variables resulted in improvements in fit across the landscape, even though annual variation in weather may have been a dominant factor during the estimation of coefficients.

8.1.4 Comparison of the Accuracy of Various Modelling Approaches

The utility of process-based modelling was clearly demonstrated by the comparisons between 3PG and growth and yield models. Results showed that empirical model accuracy can be increased further by incorporating physiological processes that describe

growing conditions on sites. Moreover the process-based model 3PG predicted basal area better than a statistically based (state-space) model CANTY even though stem number predictions were less precise. 3PG also predicted basal areas that were comparable to those predicted by models CanSPBL(water), and CanSPBL(1.2). This result shows that process-based models can give results for basal area prediction comparable to those of statistically-based models while maintaining the ability to make predictions under changing environmental conditions, such as varying climate and soil type. The model 3PG also has many physiological outputs such as estimated, carbon pools, and leaf area that can be of use to forest managers interested in carbon accounting under the Kyoto protocol. The model did not perform as well for the prediction of stocking compared to the other models tested. This effect may have been due to 3PG's use solely of the -3/2 power law to predict mortality events. The -3/2 power law would only predict mortality events due to overcrowding and would be of little use in cases of mortality caused by wind, or insect attacks. Further, basal areas output from 3-PG are determined by inverting the allometric equation that describes the relationship between individual stem mass and diameter, calculating average basal area per tree and then multiplying by stem number, so if predicted stem numbers are inaccurate, then basal area will also be affected.

The state-space model CANTY performed the worst in terms of basal area and mean top height prediction. This may have been a result of the critical dependence of state space models on site index. As the model CANTY was fitted to data of only short re-measurement intervals and lower elevation sites on the Canterbury plains, the model did not perform well predicting mean top height at elevations and at longer intervals. Site index is calculated as mean top height of the stand at age 20, so the problem with fitting mean top height can be compounded when trying to predict basal area with site indices that were originally biased.

8.2 Some Specific Views

8.2.1 Leaf Area Prediction Bias and Model Calibration

Two final models were developed to predict stand-level LAI from stand state variables. LAI estimated with these models can be used as an input in ecosystem models that simulate carbon balance and hydrological cycles. The two final models were surprisingly

simple in form, depending only on stand age, mean top height, and elevation. Results from this study demonstrate the insensitivity of LAI to stocking in older stands. Elevation in the equations seems to be a surrogate for temperature, rainfall, soil fertility, and depth of ground water.

During model fitting, there were no apparent patterns in residual plots against other possible predictor variables, such as stocking and rainfall. This pattern of simplicity has also been reported by Pinkard and Neilsen (2003), who found similar results for stand leaf areas in plantation-grown eucalypts. They reported canopy leaf areas that were nearly independent of plantation spacing, having only a moderate effect at spacings below 833 trees/ha. These results are also supported by Savill *et al.* (1997) who summarised the general effects of spacing in even aged stands as: once canopy closure has occurred and a site is being fully used in widely spaced stands, dry matter increment is similar to that in close spacings on the same site.

Canopy leaf area was found to increase on some sites at lower elevations. This effect can be viewed by increasing residual errors for both models 1 and 2 at elevations very close to zero, in coastal forests. A depth to ground water in wells in the Canterbury region is shown in Figure 6 of chapter 5. Increased growth of LAI on low altitude sites may have been due to tree roots tapping into ground water sources very close to the surface. Future research in the areas of canopy leaf area index modelling or growth modelling could focus on including this effect.

The developed models are limited to plantation-grown radiata pine. The models should be used in geographic regions that have similar climatic and soil characteristics as those within the study area, and should only be used for simulation in stands that range from 10 to 32 years of age, where canopy closure is likely to have occurred. There are discrepancies between actual LAI values, and those estimated with hemispherical photography usually due to canopy clumping and non-differentiation of foliage and non-foliage elements in the canopy (Martens *et al.,* 1993). Any use of models 1 and 2 that requires an absolute measure of leaf area will also require a calibration of model results (relative LAI) with actual values of LAI measured with destructive sampling. Mason *et al.* (in prep) have calibrated both hemispherical photographs and LAI2000 (LI-COR, 1991) measurements of LAI with destructively sampled estimates of LAI for radiata pine

growing in Canterbury. To convert LAI from models 1 and 2 to actual LAI estimated from destructive sampling Mason *et al.* (in prep) suggest a linear adjustment described by equation 8.1:

$$LAI_{ds} = LAI \cdot 3.676737832 - 1.386871774 \qquad (8.1)$$

Where LAI_{ds} = LAI obtained from destructive sampling, and LAI = Leaf area index obtained from models 1 and 2.

Mason *et al.* (in prep), found an r^2 of 0.72 when calibrating destructive sampling measurements with hemispherical photography analysed in GLA. They found no evidence the equations were affected by site factors, clones, or stocking.

The absolute leaf areas obtained after correction was made using equation 7 range from 1.1 to 7.1 with an average value of 6 for models 1 and 2. Using the Beer-Lambert equation (assuming k = 0.5 and $\beta = 0$), absolute LAI values of 6.0 or above are required to absorb at least 95% of the light. In this experiment, LAI was greater than 6.0 at stand ages greater than 15 years and, only at stockings greater than 325 stems per hectare. It is therefore concluded that in high productivity radiata pine plantations, leaf area is relatively constant regardless of spacing, provided stocking is above a threshold for canopy closure.

8.2.2 CanSPBL(water) Data Requirements and Prediction Accuracy

The model CanSPBL(water) that incorporates an index of root zone water balance over the simulation period requires inputs of stand level mensuration measurements, climatic data, and soils characteristics. These inputs are initial basal area, initial mean top height, initial stocking, initial maximum diameter, initial standard deviation of diameter, initial stand age, and final stand age. The data required to run the water balance component of the model are monthly values of maximum temperature, minimum temperature, solar radiation, vapour pressure deficit, leaf area index, rainfall, soil type, initial estimates of available soil water, maximum available soil water, and minimum available soil water. Obtaining these inputs for a given forest can be time consuming and expensive as they

require inventory measurements of initial conditions and climate and soil data that may be costly to purchase. Currently in New Zealand BIOCLIM and NEWA climate data are only available to researchers and companies willing to pay for these databases. The problems associated with acquiring this data may not warrant the 1 – 3% improvement in accuracy that these models offer. Although, improvements in availability of New Zealand weather information similar to that in the United States (e.g.: http://www.daymet.org) will facilitate the use of models that require climatic inputs and may make them more cost effective.

There is also the issue of even lower precision for the models of mean top height and standard deviation of diameter using average values of climatic inputs to describe future conditions where monthly and yearly variation may not be available. Model precision could be increased by considering the effect of drought only in months where growth is likely to occur. Drought in the winter months may not have the same effect on growth as drought that occurred during the growing season. Future research in this area could focus on trying to add a weight to ASW deficits according to the month in which they occur.

8.3 Future Research in Hybrid Modelling

This study demonstrates some of the strengths and weaknesses of various modelling approaches and takes an objective look at model accuracy from those different approaches. Future research in this area should focus on increasing model accuracy and robustness by incorporating both statistical and physiological approaches into a single model. The physiological process-based model 3PG showed promise in basal area prediction accuracy, while maintaining the ability to make predictions under changing climatic conditions but fell short in its ability to predict stocking at the end of the simulation. The hybrid approach used in CanSPBL(water) showed little improvement for the prediction of basal area, mean top height and diameter distribution over the model CanSPBL(1.2), but showed promise in the prediction of stocking using average climatic inputs. The state space model CANTY also showed promise in predicting final crop stocking. Both models CanSPBL(1.2) and CanSPBL(water) have the ability to describe diameter distributions in the modelled forest. These results indicate that some hybrid of the two approaches for a given area may produce the most accurate and robust approach to predict stand growth.

A future model may be a mix of both the statistical and physiological process-based approaches, where carbon balance and stand growth are calculated using a physiological process-based model, and diameter distribution and stocking may be regional statistical functions. Outputs from a model of this type may be more useful to managers and scientists alike as predictions of growth will be more accurate, and can be made under changing climatic conditions. A model of this type would also have the ability to output estimates of carbon pools in stems, foliage, and roots, and also be able to describe the diameter distribution of the stand.

CHAPTER 9

Summary of Conclusions

This study covers a broad research area in forest modelling systems. The main goal of the project was to evaluate a range of tree and stand-level modelling approaches, from traditional growth and yield systems to models that explicitly represent environmental influences, photosynthesis and carbon allocation, such as the 3-PG model (Landsberg and Waring, 1997). Models were fitted using data from a specific region managed by a forest company, and their performance was assessed when they were applied to an independent dataset. The study involved some biomass and canopy measurements to provide a basis for a sub-model of leaf area that was used as an input into other models. This new direction in growth and yield modelling will hopefully help bridge the gap between statistical and physiological approaches to forest modelling and give some insight as to what level of resolution is appropriate for answering questions related to forest management.

9.1 Validation of Existing Growth and Yield Model CanSPBL

The existing growth and yield model CanSPBL(1.0) was examined to determine how closely the model's behaviour fitted measurements of stands within the SPBL estate. Regression analyses and graphical procedures revealed that the models of mean top height, basal area, and stocking showed bias in projection estimates that suggested there would be benefits in updating the existing model with a current dataset that more completely covers the range in altitudes on the SPBL estate. Plots and model efficiency measures showed a general bias in under-prediction of mortality, under prediction of basal area, and over-prediction of mean top height. Residual plots for stocking showed bias in under prediction of mortality, especially for lower and higher altitude forests. This bias in stocking prediction with respect to elevation was also shown with an average model bias of -125.8 for sands, - 6.9 for plains, and -28.5 for hills plots. These same trends followed for model efficiency calculations at the three site types (Table 3.1). LOESS analysis on stocking residuals showed an increase in residuals (also indicating an under prediction in mortality) at elevations below 80m and above 250 (m) (Figure 3.13).

The basal area model showed similar trends of bias in under prediction as the stocking model (Figures 3.5 and 3.6. Areas of bias were also concentrated in plots classified as sands or hills (Table 3.1). LOESS analysis on basal area showed bias at similar elevations but at a smaller magnitude than indicated by AMB and EF measures (Figure 3.12). However, residual plots for mean top height (Figures 3.2 and 3.3) showed a bias in over prediction of height, especially for lower and higher altitude forests. This trend in mean top height was also described by model efficiency measures increasing for sands and hills forest types (Table 3.1)

9.2 Updating CanSPBL(1.0)

A new model CanSPBL(1.2) was built for use at a stand level updating equations used to fit the existing model CanSPBL(1.0). Model components of mean top height, basal area per hectare, stems per hectare, and diameter distribution were developed by using non linear least squares regression techniques to select appropriate equation forms. Explanatory variables to improve model prediction were tested and incorporated into models where appropriate. The effect of elevation was added to models of mean top height, basal area, and diameter distribution. A polymorphic Gompertz equation displayed the best fit for basal area, while a polymorphic Schumacher equation displayed the best fit for mean top height. The diameter distribution model of maximum diameter displayed the best fit with a polymorphic Weibull equation, while the standard deviation of diameter model fit best with a polymorphic Gompertz equation.

Many different modelling approaches were tried to model the number stems per hectare at the end of a simulation period. Logistic two step modelling approaches did not show an improvement over a standard 3 parameter difference equation, and the inclusion of maximum wind speed over the measurement interval resulted in only a modest improvement of 3.6 %. A system of modelling mortality with a three parameter difference equation, using a dataset filtered by a mortality severity index was chosen as the final modelling approach. This method reduced the modelling dataset to only include mortality events that were characteristic of the manager's view of average mortality. A mortality severity index based on the -3/2 power law was used as a basis to filter the

modelling data set for mortality and a final model was built with 80% of the original mortality records under the recommendation of managers at SPBL.

All model parameters were tested against an auto-correlation free dataset for significance. Tests showed that all but one parameter was significant at a 95% confidence level. The model with the insignificant parameter was kept as a final equation form after validating the model to see if in fact that form made the most accurate predictions against a validation dataset.

The entire model was validated against an independent data set and compared to a the original CanSPBL(1). CanSPBL(1,2) showed improvements in MSE of 4 to 46% after being updated with a more current and complete dataset that covered a better range of elevation and stand ages across the SPBL estate. Residual distributions for the updated model showed a better fit overall for skewness and kurtosis, with the exception of the model for mean top height where skewness is slightly higher for the updated model. The hypothesis of normality of residuals was not rejected for both models with the Kolmogorov-Smirnov test.

Final Basal Area Model

$$G_2 = e^{\left(\ln(G_1)\cdot e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)} + \frac{\left(\alpha_0+\alpha_1\cdot elev+\alpha_2\cdot elev^2+\alpha_3\cdot((elev-450)\cdot X)\right)}{10000}\cdot\left(1-e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right)\right)}$$

(9.1)

Where G_2 is the future basal area (m^2 / ha), G_1 is the initial basal area (m^2 / ha), T_1 is the initial stand age (years), T_2 is the final stand age (years), *elev* is the stand elevation (m), X is a binary indicator variable, $X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450, and β , γ, α_0, α_1, α_2, and α_3 are parameters whose values are listed in table 9.1.

Table 9.1: Parameters for basal area model (Equation 9.1), standard errors, and approximate 95% confidence limits calculated with 3666 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.1628	0.00201	0.1589	0.1668
γ	0.00261	0.00004	0.00253	0.00269
α_0	44797	255.1	44296.8	45297.2
α_1	-8.0659	1.4223	-10.8544	-5.2774
α_2	0.0491	0.00272	0.0437	0.0544
α_3	-37.8996	2.51	-42.8209	-32.9783

Final Mean Top Height Model

$$MTH_2 = e^{\ln(MTH_1)\cdot\left(\frac{T_1+\gamma}{T_2+\gamma}\right)^{\beta}+\frac{(\alpha_0+\alpha_1\cdot elev+\alpha_2\cdot elev^2+\alpha_3\cdot(elev-450)\cdot X)}{10000}\cdot\left(1-\left(\frac{T_1+\gamma}{T_2+\gamma}\right)^{\beta}\right)}$$

(eq. 9.2)

Where MTH_2 is the future mean top height (m), MTH_1 is the initial mean top height (m), T_1 is the initial stand age (years), T_2 is the final stand age (years), *elev* is the stand elevation (m), X is a binary indicator variable, $X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450, and β , γ , α_0 , α_1 , α_2 , and α_3 are parameters whose values are listed in table 9.2.

Table 9.2: Parameters for mean top height model (Equation 9.2), standard errors, and approximate 95% confidence limits calculated with 3666 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.7613	0.0466	0.6699	0.8526
γ	4.4616	0.5125	3.4568	5.4664
α_0	44521.7	609.6	43326.4	45716.9
α_1	-7.3498	1.1851	-9.6733	-5.0264
α_2	0.0342	0.00255	0.0292	0.0392
α_3	-34.604	2.8022	-40.0982	-29.1098

Final Models for Maximum Diameter and Standard Deviation of Diameter

$$DMAX_2 = DMAX_1 \cdot e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)} + \left(\alpha_0 + \alpha_1 \cdot elev^2 + \alpha_2 \cdot (elev - 450)\cdot X + \alpha_3 \cdot \frac{1}{N_1}\right)\cdot\left(1 - e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)}\right)$$

(9.3)

$$DSTD_2 = e^{\left[\ln(DSTD_1)\cdot e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right]}\cdot e^{\left[\frac{\left(\alpha_0+\alpha_1 * elev^2+\alpha_2\cdot(elev-450)\cdot X\right)}{10000}\cdot\left(1-e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right)\right]}$$

(9.4)

Where $DMAX_2$ is the future maximum diameter (m), $DMAX_1$ is the initial maximum diameter (m), $DSTD_2$ is the future standard deviation of diameter (m), $DMAX_1$ is the initial standard deviation of diameter (m),T_1 is the initial stand age (years), T_2 is the final stand age (years), *elev* is the stand elevation (m), X is a binary indicator variable, $X = 0$ if elevation < 450, N1 is the initial stocking (stems / ha) and $X = 1$ if elevation ≥ 450, and β , γ , α_0 , α_1 , α_2 , and α_3 are parameters whose values are listed in table 9.3 (for maximum diameter eq. 9.3) and table 9.4 (for standard deviation of diameter eq. 9.4).

Table 9.3: Parameter values for maximum diameter model (Equation 9.3). Also shown are standard errors, and approximate 95% confidence limits for each parameter. Statistical values are calculated with 3666 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
γ	0.3377	0.0378	0.2637	0.4118
β	0.4128	0.0333	0.3476	0.478
α_0	66.0643	3.1822	59.8252	72.3034
α_1	0.000165	0.000011	0.000144	0.000186
α_2	-0.1967	0.018	-0.232	-0.1614
α_3	8257.1	699	6886.6	9627.6

Table 9.4: Parameter values for standard deviation of diameter model (Equation 9.4). Also shown are standard errors, and approximate 95% confidence limits for each parameter. Statistical values are calculated with 3666 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.0725	0.00178	0.069	0.076
γ	0.000458	0.000061	0.000337	0.000579
α_0	24257.6	296.6	23676.1	24839.1
α_1	0.0194	0.000979	0.0175	0.0213
α_2	-33.2487	3.1668	-39.4576	-27.0398

Final Mortality Model

$$N_2 = \left(N_1^{\,c} + \frac{a}{100000} \cdot \left(T_2^{\,b} - T_1^{\,b} \right) \right)^{\left(1/c\right)} \qquad (9.5)$$

Where N_2 is the future stocking (stems / ha), N_1 is the initial stocking (stems / ha), T_1 is the initial stand age (years), T_2 is the final stand age (years).

Table 9.5: Stocking model parameters for equation 9.5, standard errors, and approximate 95% confidence limits calculated with 3280 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
a	0.0074	0.00357	0.000409	0.0144
b	2.4358	0.1414	2.1585	2.7131
c	-1.0169	0.0478	-1.1106	-0.9231

9.3 Modelling Leaf Area Index

Two models to predict canopy leaf area index in radiata pine plantations in the Canterbury region were developed. The equations estimate leaf area index as measured with hemispherical photography, and analysed with the GLA software. The equations developed were:

Model 1

$$LAI = e^{(0.8137 - 0.0008 * elevation + \frac{0.1508 * elevation^2}{100000} - \frac{751802}{age^{6.422 - 0.00101 * elevation}})} \quad (9.6)$$

Model 2

$$LAI = e^{(0.8232 + -0.00085 * elevation + \frac{0.1475 * elevation^2}{100000} - \frac{121604}{MTH^{4.9897}})} \quad (9.7)$$

Independent variables included elevation, stand age, and mean top height. The following variables were tested and were not significantly related to LAI: stocking, basal area, average diameter at breast height, and average annual rainfall. Residuals of model 1 ranged between -0.46 and 0.53, with an average residual of -0.0016 (MSE = 0.0585), while residuals of model 2 ranged between -0.42 and 0.55, with an average residual of -0.0005 (MSE = 0.0533). Testing with an independent dataset indicated that the models were more biased than expected against elevation, but overall performance of the model was similar to that found during estimation of model coefficients. Model 2 is recommended for use as it displayed a reduced overall bias and a reduction in bias with respect to independent variables when tested with an independent data set.

9.4 Incorporating an Index of Root Zone Water Balance into the Existing Growth and Yield Model CanSPBL(1.2)

A new model was built CanSPBL(water) by incorporating an index for root zone water balance over the growth interval into an existing growth and yield model CanSPBL(1.2). Model components of mean top height, basal area per hectare, stems per hectare, and diameter distribution were developed by using non linear least squares regression techniques to select appropriate equation forms. Explanatory variables to improve model prediction were tested and incorporated into models where appropriate. The effect of elevation, and ASW deficit was added to models of mean top height, basal area, and diameter distribution. While the effect of ASW deficit alone was added to the model of mortality.

A polymorphic Gompertz equation displayed the best fit for basal area, while a polymorphic Schumacher equation displayed the best fit for mean top height. The

diameter distribution model of maximum diameter displayed the best fit with a polymorphic Weibull equation, while the standard deviation of diameter model fit best with a polymorphic Gompertz equation. All model parameters were tested against an auto-correlation free dataset for significance. Tests showed that all parameters were significant at a 95% confidence level. The entire updated version of the model CanSPBL(water) was validated against an independent data set and compared to a the original CanSPBL(1.2). CanSPBL(water) showed improvements in MSE of 1 to 3% after the effects of ASW deficit were incorporated into the model, however the model of maximum diameter showed a worse fit by 0.78%. Residual distributions for CanSPBL(water) showed a worse fit overall for skewness and kurtosis. The hypothesis of normality of residuals was not rejected for both models with the Kolmogorov-Smirnov test.

To model the number of stems per hectare at the end of a simulation period a mortality severity index based on the -3/2 power law was used as a basis to filter the modelling data set. A final model was built with 80% of the original mortality records under the recommendation of managers at SPBL. Residual mean square error for the model of stocking was decreased by 2.84% by incorporating ASW deficit calculated by the equation (9.8). Residual mean square error was reduced further by 7.8% by using average climatic inputs into the model. Future stocking is dependent on the current stocking (stems / ha), initial stand age (years), final stand age (years), elevation squared (m^2), ASW deficit (mm), and ASW deficit squared (mm^2).

$$N_2 = \left(\frac{1}{\sqrt{N_1}} + \left(\alpha_0 + \alpha_1 \cdot ASW + \alpha_2 \cdot ASW^2\right) \cdot \left(\left(\frac{T_2}{100}\right)^2 - \left(\frac{T_1}{100}\right)^2 \right) \right)^{-2} \qquad (9.8)$$

Table 9.6: Stocking model parameters for (Equation 9.8), standard errors, and approximate 95% confidence limits calculated with 2226 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
α_0	0.0462	0.00313	0.0401	0.0524
α_1	-0.00035	0.000119	-0.00058	-0.00012
α_2	5.83E-06	8.13E-07	4.23E-06	7.42E-06

Residual mean square error for the model of basal area was reduced by 3.77% by incorporating ASW deficit calculated by the equation (9.9). Future basal area is dependent on the current stand basal area (m^2 / ha), initial stand age (years), final stand age (years), elevation squared (m^2), a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450), ASW deficit (mm), and ASW deficit squared (mm^2).

$$G_2 = e^{\left(\ln(G_1) e^{\left(-\beta \cdot (T_2 - T_1) + \gamma \cdot \left(T_2^2 - T_1^2 \right) \right)} + \frac{\left(\alpha_0 + \alpha_1 \cdot elev^2 + \alpha_2 \cdot ((elev - 450) \cdot X) + \alpha_3 \cdot ASW + \alpha_4 \cdot ASW^2 \right)}{10000} \cdot \left(1 - e^{\left(-\beta \cdot (T_2 - T_1) + \gamma \cdot \left(T_2^2 - T_1^2 \right) \right)} \right) \right)}$$

(9.9)

Table 9.7: Parameters for basal area model (Equation 9.9), standard errors, and approximate 95% confidence limits calculated with 3659 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.153	0.00219	0.1487	0.1573
γ	0.00249	0.000038	0.00242	0.00257
α_0	44311.5	229	43862.6	44760.4
α_1	0.0358	0.000708	0.0344	0.0372
α_2	-31.7278	2.1729	-35.9881	-27.4675
α_3	35.2148	4.363	26.6604	43.7691
α_4	-0.3759	0.0368	-0.4481	-0.3038

Residual mean square error for the model of mean top height was reduced by 2.95% by incorporating ASW deficit calculated by the equation (9.10). Future mean top height is dependent on the current mean top height (m), initial stand age (years), final stand age (years), elevation (m), elevation squared (m^2), a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450), ASW deficit (mm), and ASW deficit squared (mm^2).

$$MTH_2 = e^{\ln(MTH_1) \left(\frac{T_1 + \gamma}{T_2 + \gamma} \right)^{\beta} + \frac{\left(\alpha_0 + \alpha_1 \cdot elev^2 + \alpha_2 \cdot (elev - 450) \cdot X + \alpha_3 \cdot ASW + \alpha_4 \cdot ASW^2 \right)}{10000} \cdot \left(1 - \left(\frac{T_1 + \gamma}{T_2 + \gamma} \right)^{\beta} \right)}$$

(9.10)

Table 9.8: Parameters for mean top height model (Equation 9.10), standard errors, and approximate 95% confidence limits calculated with 3659 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.6832	0.0425	0.5998	0.7666
γ	4.1388	0.4782	3.2012	5.0764
α_0	44652.2	642.1	43393.3	45911.2
α_1	0.0238	0.00103	0.0218	0.0258
α_2	-32.8019	2.5632	-37.8276	-27.7763
α_3	40.0566	4.1723	31.8761	48.2371
α_4	-0.4519	0.039	-0.5285	-0.3754

Residual mean square error for the model of maximum diameter was increased by 0.78% by incorporating ASW deficit calculated by the equation (9.11). Future maximum diameter is dependent on the current maximum diameter (cm), initial stand age (years), final stand age (years), elevation squared (m^2), a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450), the inverse of initial stocking (ha / stems), and ASW deficit squared (mm^2).

$$DMAX_2 = DMAX_1 \cdot e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)} + \left(\alpha_0 + \alpha_1 \cdot elev^2 + \alpha_2 \cdot (elev - 450)\cdot X + \alpha_3 \cdot \frac{1}{N_1} + \alpha_4 \cdot ASW^2\right)\cdot\left(1 - e^{\left(-\beta\cdot\left(T_2^{\gamma}-T_1^{\gamma}\right)\right)}\right) \quad (9.11)$$

Table 9.9: Parameter values for maximum diameter model (Equation 9.11). Also shown are standard errors, and approximate 95% confidence limits for each parameter. Statistical values are calculated with 3659 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Conifidence Limits	
β	0.3657	0.0421	0.2832	0.4483
γ	0.3973	0.0333	0.332	0.4626
α_0	64.8266	3.1145	58.7203	70.933
α_1	0.000159	0.00001	0.000139	0.000179
α_2	-0.1755	0.0176	-0.21	-0.141
α_3	8772.2	732.1	7336.8	10207.6
α_4	0.000463	0.000091	0.000284	0.000642

Residual mean square error for the model of standard deviation of diameter was decreased by 1.07% by incorporating ASW deficit calculated by the equation (9.12). Future standard deviation of diameter is dependent on the current standard deviation of diameter (cm), initial stand age (years), final stand age (years), elevation squared (m^2), a binary indicator variable X ($X = 0$ if elevation < 450, and $X = 1$ if elevation ≥ 450), ASW deficit (mm), and ASW deficit squared (mm^2).

$$DSTD_2 = e^{\left[\ln(DSTD_1)e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right]}\cdot e^{\left[\frac{\left(\alpha_0+\alpha_1*elev^2+\alpha_2\cdot(elev-450)\cdot X+\alpha_3\cdot ASW+\alpha_4\cdot ASW^2\right)}{10000}\left(1-e^{\left(-\beta\cdot(T_2-T_1)+\gamma\cdot\left(T_2^2-T_1^2\right)\right)}\right)\right]}$$

(9.12)

Table 9.10: Parameter values for standard deviation of diameter model (Equation 9.12). Also shown are standard errors, and approximate 95% confidence limits for each parameter. Statistical values are calculated with 3659 degrees of freedom.

Parameter	Estimate	Std. Error	Approximate 95% Confidence Limits	
β	0.0715	0.00182	0.068	0.0751
γ	0.000443	0.000061	0.000323	0.000562
α_0	24777.1	358.5	24074.2	25480
α_1	0.0213	0.00109	0.0192	0.0235
α_2	-39.633	3.5957	-46.6829	-32.5831
α_3	-20.4624	6.0352	-32.2953	-8.6296
α_4	0.0935	0.0463	0.00276	0.1843

9.5 Comparison of Modelling Approaches

The study described in this thesis included an objective comparison and validation of a range of model types, with the main criterion for comparison being each model's ability to match actual historical measurements of forest growth in an independent data set. The models were examined quantitatively, assessing model behaviour with a validation data set. The procedure involved graphical displays and statistical tests. Potential correlation was detected with inspection of graphical plots of residual versus predictions and

explanatory variables. The models compared were CANTY (Goulding, 1995), CanSPBL(1.2) (Pinjuv, 2005), CanSPBL-water (Pinjuv, 2005), and 3-PG (Landsberg and Waring, 1997).

Overall, the models CanSPBL(water), and CanSPBL(1.2) performed the best in terms of basal area and mean top height prediction. Both models CanSPBL(water), and CanSPBL(1.2) showed a slightly worse fit in predictions of stocking than did the model CANTY (Table 7.5). The hybrid model 3PG showed a better fit for the prediction of basal area than the statistically based model CANTY, but showed a worse fit for the prediction of final stocking than all other models (Table 7.5). In terms of distribution of residuals, CanSPBL(1.2) had overall the lowest skewness, kurtosis, and all model parameters tested significant for normality. 3PG performed the worst on average, in terms of the distribution of residuals, and all models tested positively for the normality of residual distribution (Table 7.6).

REFERENCES

Anderson, M.C., 1966. Stand structure and light penetration. Journal of Applied Ecology. 1,41-53.

ArcView GIS. Version 3.2a. Copyright 1992-2000, Environmental Systems Research Institute, Inc.

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1998. Carbon and water fluxes in a *Pinus radiata* forest subject to soil water deficit. Australian Journal of Plant Physiology. 25, 557-570.

Arneth, A., Kelliher, F.M., McSeveny, T.M., Byers, J.N., 1999. Assessment of annual carbon exchange in a water-stressed *Pinus radiata* plantation: an analysis based on eddy covariance measurements and an integrated biophysical model. Global Change Biol. 5 (5), 531-545

Atwell, B.A., Kriedemann, P.E., Turnbull, C., 1999. Plants in action; adaptation in nature and performance in cultivation. MacMillan Education Australia Pty. Ltd. 627 Chapel Street, South Yarra 3141. 664 pages.

Ballard, R., 1971. Interrelationships between site factors and productivity of radiata pine at riverhead forest, New Zealand. Plant and Soil. 35, 371-380.

Barringer, J., Wilde, H., Willoughby, J., Burgham, S., Hewitt, A., Gibb, R., Newsome, P., Rijkse, W., 1998. Restructuring the New Zealand land resource inventory to meet the changing needs for spatial information in environmental research and management. Presented at the 10'th colloquium of the spatial information research centre, University of Otago, New Zealand. 16 – 19 November, 1998.

Battaglia, M., Sands, P.J., 1997. Modelling site productivity of *Eucalyptus globulus* in

response to climatic and site factors. Australian Journal of Plant Physiology. 24, 831-850.

Battaglia, M., Sands, P.J., 1998. Process-based forest productivity models and their application in forest management. Forest Ecology and Management. 102, 13-32.

Battaglia, M., Sands, P.J., Candy, S.G., 1999. Hybrid growth model to predict height and volume growth in young *Eucalyptus globulus* plantations. Forest Ecology and Management. 120, 193 – 201.

Becker, P., Erhart, D.W., and Smith, A.P. 1989. Analysis of forest light environments Part 1. Computerized estimation of solar radiation from hemispherical canopy photographs. Agricultural and Forest Meteorology. 44, 217-232.

Beets, P.N., Robertson, K.A., Ford-Robertson, J.B., Gordon, J., Maclaren, J.P., 1999.Description and validation of C_CHANGE: A model for simulating carbon content in managed *Pinus radiata* stands. New Zealand Journal of Forestry Science. 29(3), 409-427.

Boomsma, D.B., Hunter, I.R., 1990. Effects of water, nutrients and their interactions on tree growth, and plantation forest management practices in Australasia: a review. Forest Ecology and Management. 30, 455 – 476.

Burkhart, H.E., Tennent, R.B., 1977. Site index equations for radiata pine in New Zealand. New Zealand Journal of Forestry Science. 7(3), 408-416.

Canham, C.D., 1995. GLI/C: Software for calculation of light transmission through forest canopies using colour fisheye photography. Copyright © 1995: Institute of Ecosystem Studies, Millbrook, NY.

Causton, D.R., 1983. A biologist's basic mathematics. Edward Arnold Publishing. London, England. Pp216.

Cescatti, A., 1998. Water flux regulation in forest stands: an activity of BAHC and EUROFLUX. Annals of Forest Science. 55(1/2), 89-102.

Chapin, F.S. III, 1991. Integrated response of plants to stress. BioScience. 41, 29-36.

Chapin, F.S. III, Walter, C.H.S., Clarkson, D.T., 1988. Growth response of barley and tomato to nitrogen stress and its control by abscisic acid, water relations and photosynthesis. Planta. 173, 352-366.

Chazdon, R.L., and Field, C.B., 1987. Photographic estimation of photosynthetically active radiation: evaluation of a computerized technique. Oecologia (Berlin). 73, 525-532.

Chen, J.M., 1996. Optically-based methods for measuring seasonal variation of leaf area index in boreal conifer stands. Agricultural and Forest Meteorology. 80, 135-163.

Cherry, M., Hingston, A., Battaglia, M., Beedle, C., 1998. Calibrating the LI-COR LAI-2000 for estimating leaf area index in eucalypt plantations. Tasforests. 10, 75-82.

Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association. 74, 829-836.

Cleveland, W.S., Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. Journal of the American Statistical Association. 83, 596-610.

Clutter, J.L., 1963. Compatible growth and yield models for loblolly pine. Forest Science. 9(3), 354-371.

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H., Bailey, R.L., 1983. Timber management: a quantitative approach, New York, John Wiley and Sons. 333 pages.

Cown, D.J., 1992. New Zealand radiata pine and Douglas fir: suitability for processing. New Zealand Forest Research Institute Bullitin. 168.

Cown, D.J., McConchie, D.L., Young, G.D., 1991. Radiata pine wood properties survey. New Zealand Forest Research Institute Bullitin. 50.

Deblonde, G., Penner, M., Royer, A., 1994. Measuring leaf area index with the LI-COR LAI-2000 in pine stands. Ecology. 75(5), 1507-1511.

Dupoey, J., Leavitt, S., Choisnel, E., Jourdain, S., 1993. Modelling carbon isotope fractionation in tree rings based on effective evapotranspiration and soil water status. Plant, Cell and Environment. 16, 939-947.

Environment Canterbury, 2004. 58 Kilmore St, PO Box 345, Christchurch, New Zealand, Website: www.ecan.govt.nz.

Evans, J.R., 1989. Photosynthesis and nitrogen relationships in leaves of C_3 plants. Oecologia. 78, 9-19.

Field, C., Mooney, H.A., 1986. The photosynthesis-nitrogen relationship in wild plants. Pages 25-55 in T.J. Givinish, ed. On the Economy of plant form and function. Cambridge University Press, New York.

Fownes, J.H., Harrington, R.A., 1990. Modelling growth and optimal rotations of tropical multipurpose trees using unit leaf rate and leaf area index. Journal of Applied Ecology. 27, 886-896.

Frazer, G.W., Canham, C.D., Letzman, K.P., 1999. Gap light analyser (GLA): Imaging software to extract canopy structure and gap light transmission indices from true-colour fisheye photographs, users manual and program documentation. Copyright © 1999: Simon Fraser University, Burnaby, British Columbia, and the Institute of Ecosystem Studies, Millbrook, New York.

Garcia, O., 1984. New class of growth models for even-aged stands: *Pinus radiata* in Golden Downs Forest. New Zealand Journal of Forestry Science. 14(1), 65-88.

Garcia, O., 1988. Growth modelling – a (re)view, New Zealand Forestry. 33(3), 14-17.

Garcia, O., 1991. What is a diameter distribution? In: Proceedings of the IUFRO symposium on Integrated Management Information System, Tsubuka, Japan, October 13-16, 1991. 19 pages.

Garcia, O., 1994. The state-space approach in growth modelling. Canadian Journal of Forest Research. 24(9), 1894-1903.

Gholz, H.L., 1982. Environmental limits on aboveground net primary production, leaf area and biomass in vegetation zones of the Pacific Northwest. Ecology, 63, 469-481.

Godfrey, K., 1983. Compartmental models and their application. New York: Academic Press.

Goulding, C.J., 1995. Measurement of trees; Growth and yield models, In D. Hammond (editor): 'NZIF 1995 Forest Handbook', 1995, New Zealand Institute of Forestry (Inc.), Christchurch, P.104-107 and p.111-114.

Gower S. T., Kucharik C. J., Norman J. L. (1999) Direct and indirect estimation of leaf area index, fAPAR and net primary production for terrestrial ecosystems. Remote Sensing of the Environment. 70, 29–51.

Gower, S.T., Norman, J.M., 1991. Rapid estimation of leaf area index in conifer and broad-leaf plantations. Ecology. 75(5), 1896-1900.

Grong, P., Pu, R., Miller, J.R., 1992. Correlating leaf area index of ponderosa pine with hyperspectral CASI data. Canadian Journal of Remote Sensing. 18, 275 - 281.

Grey, D.C., 1979. Site quality prediction for *Pinus patula* in the Glengarry area, Transkei. South African Forestry Journal. 111, December 1979. 24-28.

Hale, H.E., Edwards, C., 2002. Comparison of film and digital hemispherical photography across a wide range of canopy and densities. Agricultural and Forest Meteorology. 112(1), 51-56.

Hosmer, D.W., Lemeshow, S., 1989. Applied logistic regression. Wiley, New York.

Hunter, I.R., Hunter, J,A,C., Grahaom, J.D., 1987. Pinus radiata stem volume increment and its relationship to needle mass, foliar and soil nutrients, and fertilizer inputs. New Zealand Journal of Forestry Science. 17(1), 67-75.

Jarvis, P.G., Leverenz, J.W., 1983. Productivity of temperate, deciduous and evergreen forests. Pages 233-280 *in* Land, O.L., Nobel, P.S., Osmond, C.B., Ziegler, H., editors. Ecosystem processes: mineral cycling, productivity and mans influence. Volume 12D. Physiological pant ecology: new series. Springer-Verlag, New York, New York, USA.

Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., Baret, F., 2004. Review of methods of in situ leaf area index determination; part I. theories, sensors and hemispherical photography. Agricultural and Forest Meteorology. 121, 19-35.

Johnsen, K., Samuelson, L., Teskey, R., McNulty, S., Fox, T. 2001. Process models as tools in forestry and management. Forest Science. 47, 2-8.

Kimmins, J.P., 1993. Scientific foundations for the simulation of ecosystem function and management in FORCYTE-11. Forestry Canada, Northwest region, Information Report NOR-X-328. p.88.

Kimmins, J.P., Comeau, P. G., Kurz, W., 1990. Modelling the interactions between moisture and nutrients in the control of forest growth. Forest Ecology and Management. 30, 361-379.

Kimmins, J.P., Mailly, D., Seely, B., 1999. Modelling forest ecosystem net primary production: hybrid simulation approach used in FORCAST. Ecological Modelling. 122, 195 – 224.

Korzukhin, M.D. Ter-Mikaelian, M.T., Wagner, R.G., 1996. Process versus empirical models: which approach for forest ecosystem management? Canadian Journal of Forest Research. 26, 879-887.

Kuuluvainen, T., 1991. Long-term development of needle mass, radiation interception, and stemwood production in naturally regenerated *Pinus sylvestris* stands on Empetrum-Vaccinium site type in the northern boreal zone in Finland: an analysis based on an empirical study and simulation. Forest Ecology and Management. 46, 103-122.

Land Information New Zealand, 2004. Topographic maps series 260, maps N34-31, N34-31, K38-33, and L37-33. website-www.linz.govt.nz.

Landsberg, J.J., 2003. Modelling forest ecosystems; State-of-the-art, challenges and future directions. Canadian Journal of Forest Research. 33, 385-397.

Landsberg, J.J., Gower, S.T., 1997. Applications of physiological ecology to forest management. Academic Press, San Diego, CA, p. 354.

Landsberg, J.J., Waring, R.H., 1997. A generalized model of forest productivity using simplified concepts of radiation use efficiency, carbon balance and partitioning. Forest Ecology and Management. 95, 209 – 228.

Landsberg, J.J., Coops, N.C., 1999. Modelling forest productivity across large areas and long periods. Natural Resource Modeling. 12, 383-411.

Lappi, J., 1997. A longitudinal analysis of height/diameter curves, Forest Science. 43(4), 555-570.

Leathwick, J.R. Stephens, R.T.T., 1998. Climate surfaces for New Zealand. Landcare Res. Contract Report LC9798/126. Landcare Research . Lincoln, New Zealand. 19pp.

LI-COR, 1991. Plant canopy analyser instruction manual. LI-COR Inc., Lincoln Nebraska, USA.

Livingston, N.J., Spittlehouse, D,L., 1993. Carbon isotope fractionation in tree rings in relation to the growing season water balance. In 'Stable Isotopes and Plant Carbon-Water Relations'. (Eds J.R. Ehleringer, A.E. Hall and G.D. Farquhar.) pp. 141-154. (Academic Press: San Diego.)

Loague, K., Green, R.E., 1991. Statistical and graphical methods for evaluation solute transport models: overview and application, Journal of Contaminant Hydrology., 1, 51-73.

Long, J.N., Smith, F.W., 1984. Relation between size and density in developing stands: A description of possible mechanisms. Forest Ecology and Management. 7, 191-206.

Madgwick, H.A., Jackson, D.S., Knight, P.J., 1977. Above-ground dry matter, energy, and nutrient contents of trees in an age series of *Pinus radiata* plantations. New Zealand Journal of Forestry Science. 7(3), 455-468.

Mäkelä, A., 1997. A carbon balance model of growth and self-pruning in trees based on structural relationships. Forest Science. 43, 7-24.

Mäkelä, A., Landsberg, J., Ek, A., Burk, T.E., Ter-Mikaelian, M.T., Agren, G.I., Oliver, C.D., Puttonen, P., 2000. Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiology. 20, 289-298.

Mäkelä, A., Hari, P., 1986. Stand growth model based on carbon uptake and allocation in individual trees. Ecological Modelling. 33, 205-229.

Martens, S.N., Ustin, S.L., Rousseau, R.A., 1993. Estimation of tree canopy leaf area index by gap fraction analysis. Forest Ecology and Management. 61, 91-108.

Mason, E.G., Diepstraten, M., Pinjuv, G.L., In preparation. Comparison of direct leaf area index measurements of *Pinus radiata* D.Don with estimates obtained from a Licor LAI-2000 Plant Canopy Analyser and digital analyses of hemispherical photographs. Target Journal: Agricultural and Forest Meteorology.

Mason, E.G., 1992. Decision-support systems for establishing radiata pine plantations in the central north island of New Zealand, PhD Thesis, University of Canterbury, Christchurch, New Zealand, 243 pages.

Mason, E.G., 2004. Personal Communication. Associate professor, University of Canterbury, Christchurch, New Zealand.

McLaren, R.G., Cameron, K.C., 1996. Soil science; sustainable production and environmental protection, Oxford University Press, Auckland New Zealand, Pp.60-62.

Meyers, B.J., 1988. Water stress integral-a link between short term stress and long-term growth. Tree Physiology. 4, 315-323.

Mohren, G.M.J., van Gerwen, C.P., Spitters, C.J.T., 1984. Simulation of primary production even-aged stands of Douglas fir. Forest Ecology and Management. 9, 27-49.

Monteith J. L., 1972. Solar radiation and productivity in tropical ecosystems. Journal of Applied. Ecology. 9, 747–66.

Monteith J. L., 1977. Climate and the efficiency of crop production in Britain. Phil. Trans. R. Soc. London. 281, 277 – 294.

Moore, J., 2005. Personal communication: A revised allometric model for calculating above-ground carbon using standard inventory data. Unpublished report. Forest Research, Christchurch New Zealand.

Neter, J., Wasserman, W., 1974. Applied linear statistical models: regression, analysis of variance, and experimental designs, Georgetown, Ontario, Richard D. IRWIN, INC., 842 pages.

New Zealand Forest Research Institute, 1991. FFCalc, Forestry Function Calculator user manual and tutorial, Version 1. NZFRI. Rotorua, New Zealand. FRI software series 15.

Peng, C.H., 2000a. Understanding the role of forest simulation models in sustainable forest management. Environmental Impact Assessment Review. 20, 481-501.

Peng, C.H., 2000b. Growth and yield models for uneven aged stands: past present and future. Forest Ecology and Management. 132, 259-279.

Peng, C.H., Liu, J., Dang, Q., Apps, M.J., Jiang, H., 2002. Triplex: a generic hybrid model for predicting forest growth and carbon and nitrogen dynamics. Ecological Modelling. 153, 109-130.

Pichler, M., Hager, H., Kazda, M., 2001. Contribution to the light ecology and growth of young broadleaf species (Quercus petraea, Fagus sylvatica and Acer pseudoplatanus). Centralblatt-fur-das-gesamte-Forstwesen. 118(4), 175-192.

Pienaar, L.V., Turnbull, K.J., 1973. The Chapman-Richards generalization of Von Bertalanffy's growth model for basal area growth and yield in even-aged stands. Forest Science. 19(1), 2-22.

Pierce, L.L., Running, S.W., 1988. Rapid estimation of coniferous forest leaf area index using a portable integrating radiometer. Ecology. 69(6), 1762-1767.

Pinkard, E.A., Neilsen, W.A., 2003. Crown and stand characteristics of eucalyptus nitens in response to initial spacing: implications for thinning. Forest Ecology and Management. 172, 215-227.

Raison, R.J., Khanna, P.K., Benson, M.L., Myers, B.J., McMurtrie, R.E., Lang, A.R.G., 1992. Dynamics of *Pinus radiata* foliage in relation to water and nitrogen stress: II. Needle loss and temporal changes in total foliage mass. Forest Ecology and Management. 52, 159-178.

Rich, P.M. 1990. Characterizing plant canopies with hemispherical photographs. Remote Sensing Reviews. 5(1), 13-29.

Richards, F.L., 1959. A flexible growth function for empirical use. Journal of Experimental Botany. 10, 290-300.

Richardson, B., Whitehead, D., McCracken, I.J., 2002. Root-zone water storage and growth of pinus radiata in the presence of a broom understory. New Zealand Journal of Forestry Science. 32(2), 208-220.

Running, S.W., Coughlan, J.C., 1988. A general model of forest ecosystem processes for regional applications. I Hydrologic balance, canopy gas exchange and primary production processes. Ecological Modelling. 42, 125-154.

Running, S.W., Gower, S.T., 1991. FOREST-BGC, a general model of forest ecosystem processes for regional applications. II Dynamic carbon allocation and nitrogen budgets. Tree Physiology. 9, 147-160.

Sands, P.J., Landsberg, J.J., 2002. Parameterisation of 3-PG or plantation-grown *Eucalyptus globulus*. Forest Ecology and Management. 163, 273-292.

Sands, P.J., Battaglia, M., Mummery, D., 2000. Application of a process-based models to forest management: experience with PROMOD, a simple plantation productivity model. Tree Physiology. 20(5/6), 383-392.

Savill, P., Evans, J., Aucklair, J., Falck, J., 1997. Plantation silviculture in Europe. Oxford University Press. Oxford. 143 p.

SAS Institute Inc., 2001, SAS/STAT user's guide, version 8, Cary, NC.

SPBL, 2003. Annual Report, Selwyn Plantation Board Limited. Christchurch, New Zealand. 32 pages.

Schumacher, F.X., 1939. A new growth curve and its application to timber yield studies, Journal of Forestry. 37, 819-820.

Sievänen, R., 1993. A process-based model for the dimensional growth of even-aged stands. Scandinavian Journal of Forest Research. 8, 28-48.

Sievänen, R., Burk, T., 1993. Adjusting a process-based growth model for varying site conditions through parameter estimation. Canadian Journal of Forest Research. 23, 1837-1851.

Snowdon, P., Jovanovic, T., Booth, T.H., 1999. Incorporation of indices of annual climatic variation into growth models for *Pinus radiata.* Forest Ecology and Management. 117, 187-197.

ter Steege, H, 1993. HEMIPHOT, a programme to analyse vegetation indices, light and light quality from hemispherical photographs. Copyright ©1993: The Tropenbos Foundation, Wageningen, The Netherlands.

Steinbrenner, E.C., 1963. The influence of individual soil and physiographic factors on the site index of Doulas-fir in Western Washington. Forest soil relationships in Northern America. Youngberg, C.T., (Ed.). Oregon State University Press.

Tickle, P.K., Coops, N.C., Hafner, S,D., 2001. Comparison of a forest process model (3-PG) with growth and yield models to predict productivity at Bago State Forest, NSW. Australian Forestry. 64(2), 111-122.

Vanclay, J.K., 1994. Modelling forest growth and yield: Applications to mixed tropical forests. UK: CAB International, pp. 230 -50.

Vanclay, J.K., Skovsgaard, J.P., and Hansen, C.P., 1995. Assessing the quality of permanent sample plot databases for growth modelling in forest plantations. Forest Ecology and Management. 71, 177-186.

Van Laar, A., 1967. The influence of environmental factors on the radial growth of *Pinus radiata*. South African Forestry Journal. 61, 1-16.

Von Bertalanffy, L., 1949. Problems of organic growth. Nature. 163(4135), 156-159.

Von Bertalanffy, L., 1957. Quantitative laws in metabolism and growth. Quarterly Review of Biology. 32, 217-231.

Walcroft, A.S., Silvester, W.B., Whithead, D., Kelliher, F.M., 1997. Seasonal changes in stable isotope ratios within annual rings of *Pinus radiata* reflect environmental regulation of growth processes. Australian Journal of Plant Physiology. 24, 57-68.

Waring, R.H., Schroeder, P.E., and Oren, R., 1982. Application of the pipe model theory to predict canopy leaf area. Canadian Journal of Forest Research. 12, 556-560.

Watt, M.S., Whitehead, D., Richardson, B., Mason, E.G., Leckie, A.C., 2003. Modelling the influence of weed competition on the growth of young *Pinus radiata* at a dryland site. Forest Ecology and Management. 178, 271-286.

Webb, W.L., Lauenroth, W.K., Szarek, S.R., and Kinerson, R.S., 1983. Primary production and abiotic controls in forests, grasslands, and desert ecosystems in the United States. Ecology, 64, 134-135.

Wells, J.M., 1990. Some indirect methods of estimating canopy structure. Remote Sensing Reviews. 5(1), 31-43.

West, P.W., 1995. Application of regression analysis to inventory data with measurements on successive occasions. Forest Ecology and Management. 71, 227-234.

West, P.W., Ratkowsky, D.A., Davis, A.W., 1984. Problems of hypothesis testing of regressions with multiple measurements from individual sampling units. Forest Ecology and Management. 7(3), 207-224.

Whitehead, D., Grace, J.C., Godfrey, M.J.S., 1990. Architectural distribution of foliage in individual *Pinus radiata* D. Don crowns and the effects of clumping on radiation interception. Tree Physiology. 7, 135-155.

Whitehead, D., Leathwick, J.R., Walcroft, A.S., 2001. Modeling annual carbon uptake for the indigenous forests of New Zealand. Forest Science. 47(1), 9-19.

Wilson, J.W., 1981. Analysis of growth, photosynthesis and light interception for single plants and stands. Annals of Botany. 48, 507-512.

Woollons, R.C., Haywayrd, M.J., 1985. Revision of a growth and yield model for radiata pine in New Zealand. Forest Ecology and Management. 11, 191-202.

Woollons, R.C., Whyte, A.G.D., Xu, L., 1990. The Hossfeld function: an alternative model for depiction of stand growth, Japanese Journal of Forest Environment. 15, 25-35.

Woollons, R.C., 1998. Even-aged stand mortality estimation through a two-step regression process. Forest Ecology and Management. 105, 189-195.

Woollons, R.C., Snowdon, P., Mitchell, N.D., 1997. Augmenting empirical stand projection equations with edaphic and climatic variables. Forest Ecology and Management. 98, 267-275.

Yang, R.C., Kozak, A., Smith, J.H.G., 1978. The potential of Weibull-type functions as flexible growth curves. Canadian Journal of Forest Research. 8(4), 424-431.

Yoda, K., Kira, T., Kira, H., Hozumi, H., 1963. Intraspecific competition among higher plants XI. self-thinning in overcrowded pure stands under cultivated and natural conditions. Journal of Biology, Osaka City University, Series D. 14, 107-129.

Zhang, L., Walker, G.R., Dawes, W.R. 2002. Water balance modelling: concepts and applications. In: McVicar, T.R., Li Rui, Walker, J., Fitzpatrick, R.W., Liu Changming (Eds), Regional water and soil assessment for managing sustainable agriculture in china and Australia, ACIAR Monograph No. 84, 31-47.

Zhao, W., 1999. Growth and yield modelling of *Pinus radiata* in Canterbury. New Zealand. Forestry thesis for PhD, School of Forestry, University of Canterbury. 192 pages.